Proposizione 5.6.4 (Formula di Grassmann) *Siano V e W due sottospazi di uno spazio vettoriale U. Sussiste allora la seguente relazione tra le dimensioni di V, W e $V+W$:*

$$\dim (V+W) = \dim V + \dim W - \dim (V \cap W) .$$

Dim. Sia $\mathcal{B} = \{k_1, \dots, k_h\}$ una base di $V \cap W$. In particolare $h = \dim V \cap W$. Sia $\{v_1, \dots, v_{n-h}\}$ un completamento di $\mathcal{B}$ ad una base di V, mentre sia $\{w_1, \dots, w_{m-h}\}$ un completamento di $\mathcal{B}$ ad una base di W (qui $n = \dim V$ e $m = \dim W$). Per avere la tesi basta dimostrare che $\mathcal{C} = \{k_1, \dots, k_h, v_1, \dots, v_{n-h}, w_1, \dots, w_{m-h}\}$ è una base di $V+W$.

I vettori di $\mathcal{C}$ generano $V+W$. Infatti sia $x \in V+W$. Per definizione, esistono due elementi $v \in V$ e $w \in W$ tali che $x = v + w$. Poiché $v \in V$, e $\{k_1, \dots, k_h, v_1, \dots, v_{n-h}\}$ è una base di V, si ha che:

$$v = \lambda_1 k_1 + \cdots + \lambda_h k_h + \mu_1 v_1 + \cdots + \mu_{n-h} v_{n-h} .$$

Per analoghi motivi, si ha che:

$$w = \eta_1 k_1 + \cdots + \eta_h k_h + \nu_1 w_1 + \cdots + \nu_{m-h} w_{m-h} .$$

In conclusione,

$$\begin{aligned} x &= v + w \\ &= (\lambda_1 + \eta_1) k_1 + \cdots + (\lambda_1 + \eta_1) k_h + \mu_1 v_1 + \cdots + \mu_{n-h} v_{n-h} + \\ &\quad + \nu_1 w_1 + \cdots + \nu_{m-h} w_{m-h} . \end{aligned}$$

Resta da far vedere che i vettori di $\mathcal{C}$ sono linearmente indipendenti. Sia allora

$$0 = \lambda_1 k_1 + \cdots + \lambda_1 k_h + \mu_1 v_1 + \cdots + \mu_{n-h} v_{n-h} + \nu_1 w_1 + \cdots + \nu_{m-h} w_{m-h} .$$

Poniamo $v = \mu_1 v_1 + \cdots + \mu_{n-h} v_{n-h}$. Supponiamo che sia $v \neq 0$. Il vettore v è elemento di V. Inoltre risulta che:

$$v = -(\lambda_1 k_1 + \cdots + \lambda_1 k_h + \nu_1 w_1 + \cdots + \nu_{m-h} w_{m-h})$$

e quindi $v \in W$, perché $\{k_1, \dots, k_h, w_1, \dots, w_{m-h}\}$ è una base di W. Quindi $v \in V \cap W$. Poiché per ipotesi $\mathcal{B} = \{k_1, \dots, k_h\}$ è una base di $V \cap W$, v si può scrivere come combinazione lineare dei k_i con coefficienti non tutti nulli perché, ricordiamo, $v \neq 0$. Sia $v = \eta_1 k_1 + \cdots + \eta_1 k_h$. Allora

$\eta_1 k_1 + \cdots + \eta_1 k_h - \mu_1 v_1 - \cdots - \mu_{n-h} v_{n-h} = 0$. Questo è assurdo perché $\{k_1, \ldots, k_h, v_1, \ldots, v_{n-h}\}$ è una base di V. Se invece fosse $v = 0$, poiché i v_i sono linearmente indipendenti, tutti i coefficienti μ_i devono essere nulli. Inoltre, sempre se $v = 0$, anche

$$-v = \lambda_1 k_1 + \cdots + \lambda_1 k_h + \nu_1 w_1 + \cdots + \nu_{m-h} w_{m-h} = 0\,.$$

Ma poiché $\{k_1, \ldots, k_h, w_1, \ldots, w_{m-h}\}$ è una base di W, i vettori sono indipendenti e dunque anche tutti i λ_i e i ν_i debbono essere nulli. Questo conclude la dimostrazione. □

Osserviamo in particolare che

$$\dim(V + W) \leq \dim V + \dim W$$

L'uguaglianza vale se e solo se $V \cap W = \{0\}$.

Molto spesso sarà utile considerare la somma di sottospazi che hanno intersezione banale. Per questo motivo diamo la seguente definizione.

Definizione 5.6.5 *La somma di due sottospazi V e W di uno spazio vettoriale U si dice* somma diretta *se $V \cap W = \{0\}$. La somma diretta di sottospazi si indica con il simbolo $\oplus$. In altre parole la scrittura $V \oplus W$ sta ad indicare il sottospazio $V + W$ con V e W tali che $V \cap W = \{0\}$.*

Definizione 5.6.6 *Due sottospazi V e W di uno spazio vettoriale U sono* complementari *se $V \oplus W = U$.*

Esempio 5.6.7 *$V = \{(x, y, z) \in \mathbb{R}^3 \mid z = 0\}$ e $W = \{(x, y, z) \in \mathbb{R}^3 \mid x = y = 0\}$.*

Dato $u \in \mathbb{R}^3$, $u \in V \cap W$ se e solo se $u = (0, 0, 0)$. Inoltre, $u = (x, y, z)$ si può sempre scrivere come $(x, y, z) = (x, y, 0) + (0, 0, z)$. Quindi V e W sono complementari.

Osservazione 5.6.8 *Dato un sottospazio V dello spazio vettoriale U, un complementare di V esiste sempre e in generale non è unico. Che ne esista sempre almeno uno si può vedere così. Sia $\{v_1, \ldots, v_h\}$ una base di V. Consideriamo un suo completamento ad una base di U, $w_1, \ldots, w_{m-h}$, e sia W il sottospazio generato dai w_i. È facile verificare che W è effettivamente un complementare di V.*

Proposizione 5.6.9 *Supponiamo che $U = V \oplus W$. Allora:*

1) Ogni $u \in U$ *si può scrivere in modo unico come* $u = v + w$ *con* $v \in V$ *e* $w \in W$*.*

2) Se $\mathcal{B}'$ *è una base di* V *e* $\mathcal{B}''$ *è una base di* W*, allora* $\mathcal{B} = \mathcal{B}' \cup \mathcal{B}''$ *è una base di* U*.*

3) Nella base $\mathcal{B}$ *di cui sopra, un elemento* $x \in U$ *appartiene a* V *se e solo se le ultime* k *coordinate di* x *sono* 0 *(*$k = \dim W$*);* $x \in W$ *se e solo se le prime* h *coordinate di* x *sono* 0 *(*$h = \dim V$*).*

4) $\dim V + \dim W = \dim U$*.*

Dim. 1) Che ogni $u \in U$ si possa scrivere come somma di un elemento di V con un elemento di W deriva dalla definizione di somma di sottospazi. Vediamo l'unicità di tale decomposizione. Supponiamo che $u = v_1 + w_1 = v_2 + w_2$, $v_i \in V$ e $w_i \in W$. Abbiamo $v_1 + w_1 - (v_2 + w_2) = 0$ da cui $v_1 - v_2 = w_2 - w_1$. Posto $z = v_1 - v_2 = w_2 - w_1$ si ha che $z \in V \cap W$ e dunque $z = 0$. Allora necessariamente $v_1 = v_2$ e $w_1 = w_2$.
2) Sia $\mathcal{B}' = \{b'_1, \ldots, b'_h\}$ e $\mathcal{B}'' = \{b''_1, \ldots, b''_k\}$. $\mathcal{B} = \mathcal{B}' \cup \mathcal{B}''$ è un insieme di generatori di U. Se $u \in U$, $u = v + w$ con $v \in V$ e con $w \in W$, come in 1). Allora $v = \lambda_1 b'_1 + \cdots + \lambda_h b'_h$, $w = \mu_1 b''_1 + \cdots + \mu_k b''_k$, e anche u si ottiene come combinazione di elementi di $\mathcal{B}$.

I vettori di $\mathcal{B}$ sono indipendenti. Sia $\lambda_1 b'_1 + \cdots + \lambda_h b'_h + \mu_1 b''_1 + \cdots + \mu_k b''_k = 0$. Allora $\lambda_1 b'_1 + \cdots + \lambda_h b'_h = -(\mu_1 b''_1 + \cdots + \mu_k b''_k) \in V \cap W$ è il vettore nullo. Poiché i b'_i sono indipendenti per ipotesi, essendo una base di V, tutti i coefficienti λ_i sono zero e analogamente anche tutti i μ_i sono nulli.
3) Se $x \in V$, $x = \lambda_1 b'_1 + \cdots + \lambda_h b'_h$. Pensato come elemento di U, $x = \lambda_1 b'_1 + \cdots + \lambda_h b'_h + 0 b''_1 + \cdots + 0 b''_k$. Quindi le sue ultime k coordinate nella base B di U sono nulle. Analogamente, se $x \in W$, $x = 0 b'_1 + \cdots + 0 b'_h + \mu_1 b''_1 + \cdots + \mu_k b''_k$. Risultano nulle le sue prime h coordinate.

In tutti i passaggi sopra esposti vale il viceversa.
4) Segue subito dalla 2). □

In modo ricorsivo si può definire la somma di più di due sottospazi. Dati $V_1, \ldots, V_m$, sottospazi di uno spazio vettoriale V, la loro somma è per definizione

$$V_1 + \cdots + V_m = \{v_1 + \cdots + v_m \mid v_j \in V_j, j \in 1, \ldots, m\}.$$

Questo è un sottospazio vettoriale di V. La somma $V_1 + \cdots + V_m$ si dice diretta se ogni suo elemento si scrive in modo unico come somma di vettori dei sottospazi V_j (a meno dell'ordine). In tal caso scriviamo $V_1 \oplus \cdots \oplus V_m$.

Proposizione 5.6.10 *Sia $U = V_1 + \cdots + V_m$. Allora sono fatti equivalenti:*

1) $U = V_1 \oplus \cdots \oplus V_m$ *(la somma è diretta).*

2) $V_j \cap (V_1 + \cdots + V_{j-1}) = \emptyset$, *per ogni* j.

3) Presa per ogni V_j *una base* $\mathcal{B}_j$, *allora* $\mathcal{B}_1 \cup \cdots \cup \mathcal{B}_m$ *è una base di* U.

4) $\dim U = \dim V_1 + \cdots + \dim V_m$.

5.7 Esercizi

Spazi vettoriali e basi

1) Stabilire quale dei seguenti sottoinsiemi di $\mathbb{R}^3$ è un sottospazio vettoriale:

$$\begin{aligned}
A &= \{(x,y,z) \in \mathbb{R}^3 \,|\, (x-y)^2 = 0\}\,, & (5.7)\\
B &= \{(x,y,z) \in \mathbb{R}^3 \,|\, x^2 + y^2 = 0\}\,, & (5.8)\\
C &= \{(x,y,z) \in \mathbb{R}^3 \,|\, (x+y)^2 = 0\}\,, & (5.9)\\
D &= \{(x,y,z) \in \mathbb{R}^3 \,|\, x^2 - y^2 = 0\}\,, & (5.10)\\
E &= \{(x,y,z) \in \mathbb{R}^3 \,|\, x(x-y) = 0\}\,, & (5.11)\\
F &= \{(x,y,z) \in \mathbb{R}^3 \,|\, z \in \mathbb{Z}\}\,, & (5.12)\\
G &= \{(x,y,z) \in \mathbb{R}^3 \,|\, (x-y)(x+y+z) = 0\}\,, & (5.13)\\
H &= \{(x,y,z) \in \mathbb{R}^3 \,|\, x^2 + y^2 + z^2 = 1\}\,, & (5.14)\\
I &= \{(x,y,z) \in \mathbb{R}^3 \,|\, e^{(x+y+z)} = 0\}\,, & (5.15)\\
L &= \{(x,y,z) \in \mathbb{R}^3 \,|\, e^{(x+y+z)} = 1\}\,, & (5.16)\\
M &= \{(x,y,z) \in \mathbb{R}^3 \,|\, e^{(x+y+z)} = e\}\,. & (5.17)
\end{aligned}$$

2) Trovare per quali valori del parametro $k \in \mathbb{R}$ i seguenti vettori sono una base di $\mathbb{R}^3$:

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\,, \quad v_2 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}\,, \quad v_3 = \begin{pmatrix} k \\ 0 \\ 1 \end{pmatrix}\,.$$

3) Stabilire se i seguenti vettori

$$v_1 = \begin{pmatrix} 1 \\ 5 \\ 8 \end{pmatrix}\,, \quad v_2 = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}\,, \quad v_3 = \begin{pmatrix} 2 \\ 2 \\ 5 \end{pmatrix}\,,$$

sono una base di $\mathbb{R}^3$. In tal caso trovare le coordinate di $(3, 2, 1)$ rispetto a tale base.

4) Stabilire se i seguenti vettori

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\,, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\,, \quad v_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\,,$$

sono una base di $\mathbb{R}^3$. In tal caso trovare le coordinate dei vettori della base canonica rispetto a tale base.

5) In $\mathbb{C}^2$ stabilire se i seguenti vettori

$$v_1 = \begin{pmatrix} 1+i \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 1-i \end{pmatrix},$$

sono:

1. linearmente indipendenti su $\mathbb{C}$ e su $\mathbb{R}$;
2. generano $\mathbb{C}^2$ su $\mathbb{C}$ e su $\mathbb{R}$.

6) Stabilire per quali valori dei parametri $\lambda, \mu \in \mathbb{C}$ i vettori

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ \lambda \end{pmatrix}, \quad v_2 = \begin{pmatrix} i \\ \mu \\ -2i\mu \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ i \\ 2 \end{pmatrix},$$

sono linearmente indipendenti (su $\mathbb{C}$) e una base di $\mathbb{C}^3$.

7) Mostrare che l'insiemi di vettori:

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 3 \\ 0 \\ 2 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 1 \\ 5 \\ 4 \end{pmatrix},$$

genera $\mathbb{R}^3$ ed estrarre da esso una base.

8) Mostrare che l'insieme dei seguenti vettori di $\mathbb{R}^4$:

$$v_1 = \begin{pmatrix} 1 \\ 5 \\ 0 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 9 \\ -2 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 5 \\ 1 \\ 4 \end{pmatrix},$$

è linearmente indipendente e estenderlo ad una base di $\mathbb{R}^4$.

9) Dato l'insieme di vettori di $\mathbb{R}^3$

$$v_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 5 \\ 0 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 7 \\ 1 \\ 1 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix},$$

trovare la dimensione del sottospazio che essi generano. Scelta poi una base per questo sottospazio, la si estenda ad una base di $\mathbb{R}^3$.

10) Dimostrare che

$$W = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{C}^3 \,|\, x + 2iy = iz \right\}$$

è un sottospazio vettoriale di $\mathbb{C}^3$ e calcolarne una base.

11) Stabilire per quai valori del parametro $k \in \mathbb{R}$ i seguenti vettori di $\mathbb{R}^4$

$$v_1 = \begin{pmatrix} 1 \\ 5 \\ 8 \\ k \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 5 \\ 4 \\ -2 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1 \\ 0 \\ 2 \\ 1 \end{pmatrix},$$

sono linearmente indipendenti.

Sottospazi vettoriali e formula di Grassmann

12) Dimostrare che i due insiemi

$$\mathcal{S} = \{A \in M_3(\mathbb{R}) \,|\, A = {}^tA\}, \quad \mathcal{A} = \{A \in M_3(\mathbb{R}) \,|\, A = -{}^tA\},$$

sono sottospazi vettoriali di $M_3(\mathbb{R})$ e calcolare una base per ciascuno di essi (in particolare dire quali sono le loro dimensioni). Provare infine che:

$$M_3(\mathbb{R}) = \mathcal{S} \oplus \mathcal{A}.$$

13) Dati i vettori

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 2 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 2 \\ -1 \end{pmatrix},$$

si diano delle equazioni cartesiane per il sottospazio V di $\mathbb{R}^4$ da essi generato. Si dia una base di V e la si completi ad una di $\mathbb{R}^4$.

14) Si considerino i sottospazi vettoriali V e W di $\mathbb{R}^4$ definiti dalle condizioni

$$\begin{aligned} V &: \begin{cases} 2x - y + z = 0\,, \\ z - w = 0\,, \end{cases} \\ W &: x + y + z + w = 0\,. \end{aligned}$$

Stabilire se la loro somma è diretta e, in caso contrario, si determini una base per la loro intersezione $V \cap W$. Si dica che dimensione ha $V + W$.

15) Siano

$$V = \{\, {}^t(x, y, z, w) \in \mathbb{R}^4 \mid x + y = x + z = 0\}$$

e W il sottospazio di $\mathbb{R}^4$ generato dai vettori

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

Trovare una base per $U \cap V$ e $U + V$.

16) Siano

$$V = \{\, {}^t(x, y, z, w) \in \mathbb{R}^4 \mid 2x - y + 3x - 2w = 0, 2x + y + z = 0\}$$

e W il sottospazio di $\mathbb{R}^4$ generato dai vettori

$$\begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 3 \\ 1 \\ 0 \end{pmatrix}.$$

Trovare una base per $U \cap V$ e $U + V$.

Capitolo 6

APPLICAZIONI LINEARI

6.1 Applicazioni lineari

Consideriamo due spazi vettoriali V e W sullo stesso campo base $\mathbb{K}$ ed una applicazione $f : V \to W$ fra essi.

In generale, dati due elementi di V, x e y, sommarli e poi applicare al risultato la f dà un risultato diverso dall'applicare la f a x e a y e poi sommare in W le immagini. In simboli, $f(x + y) \neq f(x)+f(y)$. Qui il simbolo $+$ indica l'operazione di somma in W, che è diversa dalla somma in V. Chiarito questo, d'ora in poi utilizzeremo per comodità lo stesso simbolo $+$, fatto questo che non produce mai confusione di sorta. Discorso analogo per il prodotto per scalare: in generale, $f(\lambda x) \neq \lambda f(x)$.

Le funzioni con queste proprietà, cioè in cui si verificano le uguaglianze, sono quindi in un qualche senso poche ma è molto importante studiarle. Introduciamo dunque la seguente definizione.

Definizione 6.1.1 *Una applicazione $f : V \to W$ fra V e W, spazi vettoriali sullo stesso campo $\mathbb{K}$, si dice* applicazione lineare *o equivalentemente* omomorfismo *se sono verificate le seguenti proprietà:*

$$
\begin{aligned}
&f(x + y) = f(x) + f(y) && \forall x, y \in V\,, \\
&f(\lambda x) = \lambda f(x) && \forall x \in V\,, \quad \forall \lambda \in \mathbb{K}\,.
\end{aligned}
$$

Se f è lineare e biiettiva, f si dice isomorfismo*.*
Una applicazione lineare $f : V \to V$ viene detta endomorfismo*.*

In altre parole, una applicazione lineare f trasforma una combinazione lineare di elementi di V nella combinazione lineare delle immagini con gli stessi coefficienti:

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)\,, \quad \forall \lambda, \mu \in \mathbb{K}, x, y \in V\,.$$

Osservazione 6.1.2 *Se $f : V \to W$ è una applicazione lineare, necessariamente deve risultare che*

$$f(0_V) = 0_W\,.$$

Infatti, $f(x) = f(x + 0_V) = f(x) + f(0_V)$. Sottraendo $f(x)$ da entrambi i membri resta verificato che $f(0_V) = 0_W$.

Esempio 6.1.3

1. *$f : V \to W$ tale che $f(v) = 0_W$ per ogni $v \in V$.*
   Infatti $f(\lambda x + \mu y) = 0_W = \lambda 0_W + \mu 0_W = \lambda f(x) + \mu f(y)$.

2. *$f : \mathbb{R}^3 \to \mathbb{R}$ tale che $f(x, y, z) = ax + by + cz$, dove $a, b, c \in \mathbb{R}$.*
   Per dimostrare che f è lineare, basterà far vedere che:

   $$f(\lambda(x_1, y_1, z_1) + \mu(x_2, y_2, z_2)) = \lambda f(x_1, y_1, z_1) + \mu f(x_2, y_2, z_2)\,.$$

   E infatti:

   $$\begin{aligned} &f(\lambda(x_1, y_1, z_1) + \mu(x_2, y_2, z_2)) = f(\lambda x_1 + \mu x_2, \lambda y_1 + \mu y_2, \lambda z_1 + \mu z_2) \\ &= \lambda a x_1 + \mu a x_2 + \lambda b y_1 + \mu b y_2 + \lambda c z_1 + \mu c z_2 \\ &= \lambda(a x_1 + b y_1 + c z_1) + \mu(a x_2 + b y_2 + c z_2) \\ &= \lambda f(x_1, y_1, z_1) + \mu f(x_2, y_2, z_2). \end{aligned}$$

   Come caso particolare, la proiezione su una coordinata è lineare. La funzione che associa ad ogni elemento di $\mathbb{R}^3$ (più in generale, di $\mathbb{R}^n$) una sua coordinata, ad es. x, si ottiene come caso particolare ponendo $a = 1, b = c = 0$.

3. $f : \mathbb{R}^2 \to \mathbb{R}^2$ *tale che* $f(x,y) = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+2y \\ x \end{pmatrix}$.

$$\begin{aligned} f(\lambda x_1 + \mu x_2, \lambda y_1 + \mu y_2) &= \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda x_1 + \mu x_2 \\ \lambda y_1 + \mu y_2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \left\{ \lambda \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \mu \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \right\} \\ &= \lambda \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \mu \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \\ &= \lambda f(x_1, y_1) + \mu f(x_2, y_2). \end{aligned}$$

4. $\mathcal{F}_c : \mathcal{F}(A, \mathbb{R}) \to \mathbb{R}$, *dove* $\mathcal{F}(A, \mathbb{R})$ *è lo spazio delle funzioni definite su* A *a valori in* $\mathbb{R}$ *e* $c \in A$, $\mathcal{F}_c(f) = f(c)$. *Ossia,* $\mathcal{F}_c$ *associa all'elemento* $f \in \mathcal{F}(A, \mathbb{R})$ *il valore che* f *assume nel punto fissato* c *di* A.

5. $\mathcal{D} : \mathbb{R}[x] \to \mathbb{R}[x]$, *dove* $\mathcal{D}(p(x))$ *è la derivata del polinomio* $p(x)$*:*

$$\mathcal{D}(a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + a_1 .$$

6. $\mathrm{tr} : \mathbf{M}_n(\mathbb{R}) \to \mathbb{R}$, *dove* $\mathbf{M}_n(\mathbb{R})$ *è lo spazio delle matrici quadrate di ordine* n *e* tr *è la funzione che associa alla matrice* $A = (a_{ij})$ *la sua traccia,* $\mathrm{tr}(A) = \Sigma_{i=1}^n a_{ii}$.

Invece *non* sono applicazioni lineari le seguenti:

1. $f : \mathbb{R} \to \mathbb{R}$ tale che $f(x) = x^2$. Infatti, $f(\lambda x) = \lambda^2 x^2 = \lambda^2 f(x) \neq \lambda f(x)$.

2. $\det : \mathbf{M}_n(\mathbb{R}) \to \mathbb{R}$, dove $\det$ è la funzione che associa alla matrice A il suo determinante, con $n > 1$. Basta un controesempio.
   Se $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ e $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ si ha che:
   $\det A = 0$, $\det B = 0$, ma $\det(A+B) = \det I = 1 \neq \det A + \det B$.

Un discorso a parte va fatto per il seguente esempio.

Esempio 6.1.4 *Il coniugio su* $\mathbb{C}$. *Se* $z \in \mathbb{C}$, $z = a + ib$, *allora* $\bar{z} = a - ib$. *Ricordiamo che* $\overline{z+w} = \bar{z} + \bar{w}$, *ma che* $\overline{zw} = \bar{z}\bar{w}$. *Questo ci dice che il coniugio non è una applicazione lineare se considero come* $\mathbb{C}$ *come spazio*

vettoriale su $\mathbb{C}$. Se però considero $\mathbb{C}$ come spazio vettoriale su $\mathbb{R}$, ossia se mi limito a considerare solo la moltiplicazione per un numero reale λ, allora: $\overline{\lambda z} = \bar{\lambda}\bar{z} = \lambda\bar{z}$. (Ricordiamo che $\lambda = \bar{\lambda}$ se e solo se $\lambda \in \mathbb{R}$). Quindi il coniugio su $\mathbb{C}$ è una applicazione $\mathbb{R}$-lineare. Dunque è importante tenere presente il campo base.

È molto semplice eseguire le verifiche che dimostrano le seguenti proposizioni, che, pertanto, vengono lasciate per esercizio.

Proposizione 6.1.5 *Siano $f : V \to W$ e $g : W \to U$ due applicazioni lineari. Allora la funzione composta $g \circ f : V \to U$ è anch'essa lineare.*

Proposizione 6.1.6 *Sia $f : V \to W$ una applicazione lineare invertibile. Allora la funzione inversa $f^{-1} : W \to V$ è anch'essa lineare.*

Proposizione 6.1.7 *Siano V e W due spazi vettoriali. Indichiamo con* $\mathrm{Hom}(V, W)$ *l'insieme di tutte le applicazioni lineari f da V in W. Allora* $\mathrm{Hom}(V, W)$ *è uno spazio vettoriale.*

I seguenti risultati ci mostrano come l'essere lineare riduce di molto le possibilità di scelta, quando si debbano determinare delle applicazioni con determinate altre caratteristiche.

Lemma 6.1.8 *Siano V e W due spazi vettoriali e $\mathcal{B} = \{b_1, \ldots, b_n\}$ una base di V. Inoltre siano $f, g \in \mathrm{Hom}(V, W)$. Se, per ogni indice i, accade che $f(b_i) = g(b_i)$, allora $f = g$.*

Dim. Sia $v \in V$ un generico elemento di V. Il vettore v si potrà dunque scrivere come combinazione degli elementi della base B. $v = \lambda_1 b_1 + \cdots + \lambda_n b_n$. Verifichiamo che $f(v) = g(v)$.

$$\begin{aligned} f(v) &= f(\lambda_1 b_1 + \cdots + \lambda_n b_n) = \lambda_1 f(b_1) + \cdots + \lambda_n f(b_n) \\ &= \lambda_1 g(b_1) + \cdots + \lambda_n g(b_n) = g(\lambda_1 b_1 + \cdots + \lambda_n b_n) = g(v)\,. \end{aligned}$$

□

Teorema 6.1.9 *Siano V e W due spazi vettoriali e $\mathcal{B} = \{b_1, \ldots, b_n\}$ una base di V. Fissati n elementi qualsiasi di W, $w_1, \ldots, w_n$, esiste una unica applicazione lineare $f : V \to W$, tale che $f(b_i) = w_i$.*

Dim. L'unicità di una tale applicazione lineare discende immediatamente dal Lemma precedente. Dimostriamo la sua esistenza costruendola esplicitamente. Definiamo dunque $f : V \to W$ come segue.
Se $v \in V$ è tale che $v = \lambda_1 b_1 + \cdots + \lambda_n b_n$, allora $f(v) = \lambda_1 w_1 + \cdots + \lambda_n w_n$.
L'applicazione f è ben definita. Verifichiamone la linearità.
Se $u \in V$, $u = \mu_1 b_1 + \cdots + \mu_n b_n$, allora $u + v = (\lambda_1 + \mu_1) b_1 + \cdots + (\lambda_n + \mu_n) b_n$ e

$$\begin{aligned} f(v+u) &= (\lambda_1 + \mu_1) w_1 + \cdots + (\lambda_n + \mu_n) w_n \\ &= (\lambda_1 w_1 + \cdots + \lambda_n w_n) + (\mu_1 w_1 + \cdots + \mu_n w_n) = f(v) + f(u)\,, \\ f(\lambda v) &= \lambda\lambda_1 w_1 + \cdots + \lambda\lambda_n w_n = \lambda(\lambda_1 w_1 + \cdots + \lambda_n w_n) = \lambda f(v)\,. \end{aligned}$$

□

Questo risultato ci garantisce che possiamo assegnare a piacere le immagini degli elementi di una base, ma poi, per linearità, risulta fissata l'immagine di ogni altro elemento di V.

Proposizione 6.1.10 *Sia V uno spazio vettoriale sul campo $\mathbb{K}$, di dimensione finita. L'insieme* $\mathrm{Hom}(V, \mathbb{K})$ *delle applicazione lineari da V in $\mathbb{K}$ è uno spazio vettoriale della stessa dimensione di V.*
Lo spazio $\mathrm{Hom}(V, \mathbb{K})$ *si indica anche con V^* e viene detto spazio* duale *di V.*

Dim. Sia $\mathcal{B} = \{b_1, \ldots, b_n\}$ una base di V. Sia b_i^* l'elemento di V^* tale che $b_i^*(b_j) = \delta_{ij}$. In altre parole, b_i^* è l'omomorfismo che vale 1 sull'elemento b_i della base, e vale zero se applicato a tutti gli altri elementi della base scelta.
Dimostriamo che l'insieme $\mathcal{B}^* = \{b_1^*, \ldots, b_n^*\}$ è una base di V^*.
a) Gli elementi $b_1^*, \ldots, b_n^*$ sono linearmente indipendenti. Infatti sia $\lambda_1 b_1^* + \cdots + \lambda_n b_n^* = \underline{0}$. (Qui $\underline{0}$ indica l'elemento neutro di V^*, cioè l'omomorfismo che vale identicamente zero su V). Se fosse per assurdo $\lambda_i \neq 0$ risulterebbe:

$$\underline{0}(b_i) = 0 = (\lambda_1 b_1^* + \cdots + \lambda_n b_n^*)(b_i) = \lambda_1 b_1^*(b_i) + \cdots + \lambda_n b_n^*(b_i) = \lambda_i b_i^*(b_i) = \lambda_i$$

essendo tutti gli altri termini nulli, per la definizione di b_i^*.
b) Gli elementi $b_1^*, \ldots, b_n^*$ sono dei generatori di V^*. Infatti, sia $f \in V^*$. Poniamo, inoltre $k_i = f(b_i)$, cioè, k_i è l'elemento del campo $\mathbb{K}$ che l'omomorfismo f associa a b_i. Allora risulterà che $f = k_1 b_1^* + \cdots + k_n b_n^*$. Questo è immediato, perché $f(b_i) = (k_1 b_1^* + \cdots + k_n b_n^*)(b_i) = k_i$. Quindi i due omomorfismi, assumendo gli stessi valori sugli elementi della base di V, coincidono per il Lemma 6.1.8. □

6.2 Nucleo e immagine di una applicazione lineare

Vediamo di indagare più dettagliatamente sulla struttura e sulle proprietà delle applicazioni lineari.

Definizione 6.2.1 *Sia $f : V \to W$ un'applicazione lineare. L'insieme:*

$$f^{-1}(0_W) = \{v \in V \mid f(v) = 0_W\}$$

è detto nucleo *di f.*

Proposizione 6.2.2 *Il nucleo di un'applicazione lineare $f : V \to W$ è un sottospazio vettoriale di V. Indichiamo questo sottospazio con* $\mathrm{Ker} f$.

Dim. In primis, $\mathrm{Ker} f$ non è mai vuoto; esso contiene sempre almeno 0_V. Basta ora verificare che, se $u, v \in \mathrm{Ker} f$, allora ogni loro combinazione lineare $\lambda u + \mu v \in \mathrm{Ker} f$, cioè che $f(\lambda u + \mu v) = 0_W$. Questo è immediato, infatti $f(\lambda u + \mu v) = \lambda f(u) + \mu f(v) = 0_W + 0_W = 0_W$. □

Proposizione 6.2.3 *Sia f una applicazione lineare. Allora f è iniettiva se e solo se* $\mathrm{Ker} f = \{0_V\}$.

Dim. Supponiamo che f sia iniettiva, cioè $f(x) \neq f(y)$ se $x \neq y$. Se $x \neq 0_V$, allora $f(x) \neq f(0_V)$, ossia $f(x) \neq 0_W$. Quindi ogni x diverso da 0_V non appartiene a $\mathrm{Ker} f$.

Viceversa, supponiamo che $\mathrm{Ker} f = \{0_V\}$ e deduciamone che f è iniettiva. Supponiamo che $f(x) = f(y)$. Poiché f è lineare, si può scrivere che: $0_W = f(x) - f(y) = f(x - y) = 0$. Ciò significa che $x - y \in \mathrm{Ker} f$ e dunque, per ipotesi, $x - y = 0_V$, ossia $x = y$. □

Definizione 6.2.4 *Sia $f : V \to W$ un'applicazione lineare. L'insieme:*

$$f(V) = \{w \in W \mid \exists v \in V \textit{tale che } w = f(v)\}$$

si dice immagine *di f.*

Proposizione 6.2.5 *L'immagine di un'applicazione lineare $f : V \to W$ è un sottospazio vettoriale di W. Indichiamo questo sottospazio con* $\mathrm{Im} f$.

Dim. $0_W \in \mathrm{Im} f$ perché, per l'Osservazione 6.1.2, $f(0_V) = 0_W$. Supponiamo che i vettori w_1 e w_2 di W stiano in $\mathrm{Im} f$ e siano v_1 e v_2 le rispettive controimmagini in V, cioè $w_i = f(v_i)$, per $i = 1, 2$. Abbiamo:

$$\lambda w_1 + \mu w_2 = \lambda f(v_1) + \mu f(v_2) \, .$$

Per la linearità di f, $\lambda w_1 + \mu w_2 = f(\lambda v_1 + \mu v_2)$. Quindi, anche $\lambda w_1 + \mu w_2 \in \mathrm{Im} f$. □

Corollario 6.2.6 *Se $f : V \to W$ è lineare e U è un sottospazio di V, allora $f(U)$ è un sottospazio di* $\mathrm{Im} f$ *e quindi anche di W.*

Osservazione 6.2.7 *Se $b_1, \ldots, b_n$ sono dei generatori di V, allora i vettori $f(b_1), \ldots, f(b_n)$ sono dei generatori di* $\mathrm{Im} f$. *Infatti, sia $w \in \mathrm{Im} f$ e $v \in V$ tale che $w = f(v)$. Se $v = \eta_1 b_1 + \cdots + \eta_n b_n$, allora:*

$$w = f(v) = f(\eta_1 b_1 + \cdots + \eta_n b_n) = \eta_1 f(b_1) + \cdots + \eta_n f(b_n) \, .$$

Quindi w si scrive come combinazione dei vettori $f(b_i)$.

A questo punto verrebbe spontaneo chiedersi se gli $f(b_i)$ non costituiscano una base di $\mathrm{Im} f$. La risposta in generale è no, come c'era da aspettarsi, in quanto sappiamo che le immagini dei vettori della base possono essere scelte arbitrariamente e quindi anche non indipendenti.

Se però $v_1, \ldots, v_n$ sono dipendenti, allora anche le immagini lo sono. Infatti, se $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0_V$ con qualche coefficiente non nullo, allora:

$$f(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n) = f(0_V) = 0_W \, ,$$

quindi gli $f(v_i)$ sono dipendenti, perché non tutti i λ_i sono nulli.

Tutto questo ci permette di poter dire che:

1. *Le immagini di vettori indipendenti possono essere anche dipendenti.*

2. *Le immagini di vettori dipendenti sono necessariamente dipendenti.*

Abbiamo inoltre:

Proposizione 6.2.8 *Se $f : V \to W$ è una applicazione lineare iniettiva, allora le immagini di vettori indipendenti sono indipendenti.*

Dim. Siano $v_1, \ldots, v_h$ vettori indipendenti di V e supponiamo per assurdo che le loro immagini siano linearmente dipendenti. Sia $\eta_1 f(v_1) + \cdots + \eta_h f(v_h) = 0$ con coefficienti non tutti nulli. Allora $f(\eta_1 v_1 + \cdots + \eta_h v_h) = 0$ e quindi $\eta_1 v_1 + \cdots + \eta_h v_h \in \mathrm{Ker} f$. Ma poiché i vettori $v_1, \ldots, v_h$ sono indipendenti e i coefficienti η_i non sono tutti nulli, allora $\eta_1 v_1 + \cdots + \eta_h v_h \neq 0$; quindi, per la Propoposizione 6.2.3, la f non è iniettiva. □

L'affermazione più significativa che si può fare su questo argomento è contenuta nel seguente enunciato.

Teorema 6.2.9 *Sia $f : V \to W$ una applicazione lineare tra due spazi vettoriali V e W. Allora:*

$$\dim \mathrm{Ker} f + \dim \mathrm{Im} f = \dim V \,.$$

Dim. Sia $\mathcal{B} = \{b_1, \ldots, b_n\}$ una base di V. Allora $f(b_1), \ldots, f(b_n)$ sono dei generatori di $\mathrm{Im} f$, per Osservazione 6.2.7.
Se $\mathrm{Ker} f = \{0\}$, allora essi sono anche indipendenti, sono perciò una base di $\mathrm{Im} f$ e l'uguaglianza è dimostrata.
Se $\mathrm{Ker} f \neq \{0\}$, sia $\{k_1, \ldots, k_h\}$ una base di $\mathrm{Ker} f$. Completiamola ad una base di V con i vettori $b_1, \ldots, b_{n-h}$. Se dimostriamo che $\{f(b_1), \ldots, f(b_{n-h})\}$ è una base di $\mathrm{Im} f$, allora abbiamo vinto.
a) $f(b_1), \ldots, f(b_{n-h})$ sono indipendenti.

Sia per assurdo $\eta_1 f(b_1) + \cdots + \eta_{n-h} f(b_{n-h}) = 0$, una loro combinazione che dà lo zero di W ma con coefficienti non tutti nulli. Allora il vettore $\eta_1 b_1 + \cdots + \eta_{n-h} b_{n-h}$ di V è un elemento non nullo che sta nel $\mathrm{Ker} f$. Esso può essere dunque espresso da una combinazione dei vettori $k_1, \ldots, k_h$ che sono una base di $\mathrm{Ker} f$:

$$\eta_1 b_1 + \cdots + \eta_{n-h} b_{n-h} = \lambda_1 k_1 + \cdots + \lambda_h k_n \,.$$

Ma allora si ha:

$$\eta_1 b_1 + \cdots + \eta_{n-h} b_{n-h} - \lambda_1 k_1 - \cdots - \lambda_h k_n = 0 \,.$$

Ma questo è assurdo, perché questa combinazione ha i coefficienti non tutti nulli ma $\{k_1, \cdots, k_h, b_1, \cdots, b_{n-h}\}$ è una base di V.
b) $f(b_1), \cdots, f(b_{n-h})$ sono dei generatori di $\mathrm{Im} f$. Sia $w \in \mathrm{Im} f$ e sia $v \in V$ tale che $w = f(v)$. Poiché $\{k_1, \ldots, k_h, b_1, \ldots, b_{n-h}\}$ è una base di V,

$$\begin{aligned} v &= \eta_1 b_1 + \cdots + \eta_{n-h} b_{n-h} + \lambda_1 k_1 + \cdots + \lambda_n k_n \,, \\ w &= f(v) = f(\eta_1 b_1 + \cdots + \eta_{n-h} b_{n-h} + \lambda_1 k_1 + \cdots + \lambda_h k_n) \\ &= \eta_1 f(b_1) + \cdots + \eta_{n-h} f(b_{n-h}) \end{aligned}$$

perché $f(\lambda_1 k_1 + \cdots + \lambda_n k_n) = 0$ in quanto $\lambda_1 k_1 + \cdots + \lambda_n k_n$ sta in $\mathrm{Ker} f$. Quindi $w \in \mathrm{Im} f$ si può scrivere sempre come combinazione degli $f(b_i)$. □

Corollario 6.2.10 *Due spazi vettoriali V e W sono isomorfi se e solo se hanno la stessa dimensione.*

Dim. Sia $f : V \to W$ un isomorfismo tra gli spazi vettoriali V e W. Poiché f è suriettiva $\mathrm{Im} f = W$ e poiché f è iniettiva, $\dim \mathrm{Ker} f = 0$. Quindi $\dim V = \dim \mathrm{Ker} f + \dim \mathrm{Im} f = \dim W$.

Viceversa, supponiamo che $\dim V = \dim W = n$. Ogni base di V ha lo stesso numero di elementi di una base di W. Sia $\mathcal{B} = \{b_1, \ldots, b_n\}$ una base fissata di V e $\mathcal{B}' = \{b'_1, \ldots, b'_n\}$ una di W. Sia $f : V \to W$ l'applicazione lineare tale che $f(b_i) = b'_i$, che esiste per il Teorema 6.1.9. Quindi $W = \mathrm{Im} f$ perché sono entrambi generati dallo stesso insieme di vettori $f(b_i) = b'_i$, dunque f è suriettiva. Inoltre, $\dim \mathrm{Ker} f = \dim V - \dim \mathrm{Im} f = 0$ per cui f è anche iniettiva. La f è dunque un isomorfismo. □

6.3 Applicazioni lineari e matrici

Ricordiamo che, fissata una base (ordinata) in uno spazio vettoriale V, a ciascun vettore v di V era assegnata una n-upla di numeri reali che erano le coordinate di v e che, per l'Osservazione 5.5.2, le coordinate della somma di vettori sono la somma delle coordinate degli addendi e che le coordinate del prodotto di un vettore per uno scalare risultano essere il prodotto delle coordinate del vettore per lo scalare dato. Nel linguaggio delle applicazioni lineari, ciò si può compendiare dicendo che, fissata una base $\mathcal{B}$, esiste un omomorfismo $c_{\mathcal{B}} : V \to \mathbb{R}^n$ tale che $c_{\mathcal{B}}(v)$ sono le coordinate di $v \in V$.

In particolare, $c_{\mathcal{B}}$ risulta essere biiettiva cioè un isomorfismo, che ovviamente dipende dalla scelta della base $\mathcal{B}$ e dal suo ordinamento.

Osservazione 6.3.1 *Data una matrice A di tipo $m \times n$, risulta determinata una applicazione $f_A : \mathbb{R}^n \to \mathbb{R}^m$ nel modo seguente:*

$$\text{se } X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \text{ allora } f_A(X) = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

dove AX è il prodotto riga per colonna tra matrici (quindi appartiene a R^m). Dalle proprietà del prodotto tra matrici è facile verificare che f_A è lineare.

Esempio 6.3.2 *Sia* $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix}$. *Allora* $f_A : \mathbb{R}^3 \to \mathbb{R}^2$ *è così definita: se* $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$, $f_A(X) = AX = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 + x_3 \\ x_1 - x_2 \end{pmatrix}$.

Possiamo estendere questo procedimento di costruzione di applicazioni lineari a qualsiasi spazio vettoriale.

Osservazione 6.3.3 *Siano V e W due spazi vettoriali di dimensione rispettivamente n e m, su cui abbiamo fissato delle basi ordinate e quindi delle coordinate. Sia A una matrice $m \times n$ a valori reali. Possiamo costruire una applicazione $f : V \to W$ nel seguente modo:*
di ogni vettore $v \in V$ consideriamo la sua n-upla di coordinate X scritta come una colonna $n \times 1$ e facciamo il prodotto AX. Il risultato è un vettore colonna $m \times 1$ che individua un vettore $w \in W$, nel sistema di coordinate scelto. Si sceglie quindi $w = f(v)$ come definizione della f. Con queste scelte, f risulta lineare. Anche in questo caso scriviamo f_A per indicare la applicazione lineare costruita a partire da A. L'applicazione f_A dipende fortemente anche dalla scelta delle basi in V e W. Cambiando tali basi, cambiano le coordinate dei vettori e dunque la stessa matrice A farà corrispondere vettori diversi. Studieremo in seguito questa importante dipendenza.

Di capitale importanza risulta essere il fatto seguente che rappresenta il viceversa della precedente osservazione. In questa abbiamo visto che, fissate delle basi in V e W, ad ogni matrice di tipo opportuno risulta associata una applicazione lineare tra i due spazi. Adesso vogliamo far vedere che ad ogni applicazione lineare f corrisponde una matrice A tale che $f = f_A$.

Definizione 6.3.4 *Sia $f : V \to W$ una applicazione lineare e siano $\mathcal{B} = \{b_1, \dots, b_n\}$ e $\mathcal{B}' = \{b'_1, \dots, b'_m\}$ delle basi rispettivamente di V e di W. Definiamo A_f come la matrice che ha come i-esima colonna le coordinate (rispetto alla base $\mathcal{B}'$) di $f(b_i)$, l'immagine secondo f dell'i-esimo elemento della base $\mathcal{B}$ di V.*

A_f è una matrice $m \times n$, cioè che ha tante righe quant'è la dimensione dello spazio di arrivo W e tanta colonne quant'è la dimensione dello spazio di partenza V.

Proposizione 6.3.5 *Sia $f : V \to W$ una applicazione lineare tra gli spazi vettoriali V e W sui quali sono state fissate delle basi e sia A_f la matrice associata a f nelle basi date, secondo la costruzione precedente.*

Se le coordinate di $v \in V$ sono la n-upla $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, allora le coordinate di $f(v) \in W$ sono la m-upla $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A_f X$.

Dim. La dimostrazione è una semplice verifica. Sia $\mathcal{B} = \{b_1, \cdots, b_n\}$ la base scelta di V e $\mathcal{B}' = \{b'_1, \ldots, b'_m\}$ quella di W. Se $A_f = (a_{ij})$, per come è definita A_f, si ha che :

$$\begin{aligned} f(b_1) &= a_{11}b'_1 + \cdots + a_{m1}b'_m \,, \\ &\vdots \\ f(b_n) &= a_{1n}b'_1 + \cdots + a_{mn}b'm \,. \end{aligned}$$

Osserviamo che a_{ij} è il coefficiente di b'_i nella espressione di $f(b_j)$. Sia $v \in V$ tale che $v = \lambda_1 b_1 + \cdots + \lambda_n b_n$.

Per la linearità di f, $f(v) = f(\lambda_1 b_1 + \cdots + \lambda_n b_n) = \lambda_1 f(b_1) + \cdots + \lambda_n f(b_n)$. Quindi

$$\begin{aligned} f(v) &= \lambda_1(a_{11}b'_1 + \cdots + a_{m1}b'_m) + \cdots + \lambda_n(a_{1n}b'_1 + \cdots + a_{mn}b'_m) \\ &= (\lambda_1 a_{11} + \cdots + \lambda_n a_{1n})b'_1 + \cdots + (\lambda_1 a_{m1} + \cdots + \lambda_n a_{mn})b'_m. \end{aligned}$$

Quindi $f(v)$ ha come coordinate la m-upla

$$Y = \begin{pmatrix} \lambda_1 a_{11} + \cdots + \lambda_n a_{1n} \\ \vdots \\ \lambda_1 a_{m1} + \cdots + \lambda_n a_{mn} \end{pmatrix}.$$

È facile verificare che

$$Y = A_f \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

□

Esempio 6.3.6 *1)* $\mathcal{D} : \mathbb{R}_3[x] \to \mathbb{R}_3[x]$, *dove* $\mathcal{D}(p(x))$ *è la derivata del polinomio* $p(x) \in \mathbb{R}_3[x]$, *lo spazio dei polinomi a coefficienti reali di grado minore o uguale a 3.*

$$\mathcal{D}(a_3x^3 + a_2x^2 + a_1x + a_0) = 3a_3x^2 + 2a_2x + a_1 .$$

Consideriamo la seguente base di $\mathbb{R}_3[x]$, $\{1, x, x^2, x^3\}$. *Abbiamo:*
$\mathcal{D}(1) = 0$. *In coordinate,* $(0, 0, 0, 0)$.
$\mathcal{D}(x) = 1$. *In coordinate,* $(1, 0, 0, 0)$.
$\mathcal{D}(x_2) = 2x$. *In coordinate,* $(0, 2, 0, 0)$.
$\mathcal{D}(x_3) = 3x^2$. *In coordinate,* $(0, 0, 3, 0)$.
Quindi la matrice $A_{\mathcal{D}}$ *associata, nella base scelta, all'applicazione* $\mathcal{D}$, *è:*

$$A_{\mathcal{D}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Sia $p(x) = a_3x^3 + a_2x^2 + a_1x + a_0$ *il generico elemento di* $\mathbb{R}_3[x]$. *Le sue coordinate sono:* $X = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}$. *Allora* $\mathcal{D}(p(x))$ *ha come coordinate* $A_{\mathcal{D}}X$.
Abbiamo

$$A_{\mathcal{D}}X = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ 2a_2 \\ 3a_3 \\ 0 \end{pmatrix}$$

e $\mathcal{D}(p(x)) = 3a_3x_2 + 2a_2x + a_1$.
2) Sia $f : \mathbb{R}^4 \to \mathbb{R}^2$ *così definita:* $f(x, y, z, w) = (3x - 2z + w, -x + y - z + w)$. *Supponendo di aver fissato su entrambi gli spazi la base canonica, avremo che:*

$$\begin{aligned} f(1, 0, 0, 0) &= (3, -1) , \\ f(0, 1, 0, 0) &= (0, 1) , \\ f(0, 0, 1, 0) &= (-2, -1) , \\ f(0, 0, 0, 1) &= (1, 1) . \end{aligned}$$

La matrice A *si otterrà scrivendo in colonna le coordinate dei vettori immagine, quindi sarà:*

$$A = \begin{pmatrix} 3 & 0 & -2 & 1 \\ -1 & 1 & -1 & 1 \end{pmatrix} .$$

Abbiamo così stabilito una corrispondenza tra lo spazio $\mathrm{Hom}(V,W)$ e lo spazio $\mathbf{M}_{m,n}(\mathbb{R})$, associando ad ogni applicazione lineare $f : V \to W$, la matrice A_f, come si è fatto nella Definizione 6.3.4. Ebbene questa corrispondenza è un *isomorfismo* tra i due spazi vettoriali.

Proposizione 6.3.7 *Siano V e W due spazi $\mathbb{R}-$vettoriali di dimensione rispettivamente n e m. Sia $\mathcal{B} = \{b_1, \dots, b_n\}$ una base fissata di V e $\mathcal{B}' = \{b'_1, \dots, b'_m\}$ una di W.*
Sia $\Phi : \mathrm{Hom}(V,W) \to \mathbf{M}_{m,n}(\mathbb{R})$ l'applicazione tale che $\Phi(f) = A_f$, dove A_f è come nella Definizione 6.3.4. L'applicazione Φ è lineare invertibile.

Dim. Dimostriamo, nell'ordine, la linearità, la suriettività e la iniettività della Φ.

1) Φ è lineare, cioè

$$\Phi(\lambda f + \mu g) = \lambda \Phi(f) + \mu \Phi(g) \qquad \forall f, g \in \mathrm{Hom}(V,W)\,, \quad \forall \lambda, \mu \in \mathbb{R}\,.$$

$\Phi(\lambda f + \mu g)$ è la matrice che ha al posto (i,j) il coefficiente di b'_i del vettore $(\lambda f + \mu g)(b_j)$. Poiché f e g sono lineari, il vettore $(\lambda f + \mu g)(b_j)$ è uguale al vettore $\lambda f(b_j) + \mu g(b_j)$ e il coefficiente di b'_i nell'espressione di $(\lambda f + \mu g)(b_j)$ è uguale a λ volte il coefficiente di b'_i nell'espressione di $f(b_j)$ più μ volte il coefficiente di b'_i nell'espressione di $g(b_j)$. Quindi, essendo ciò vero per ogni i e j, risulta che

$$A_{\lambda f + \mu g} = \lambda A_f + \mu A_g\,.$$

2) Φ è suriettiva.
Sia $A \in \mathbf{M}_{m,n}(\mathbb{R})$ e sia $f_A \in \mathrm{Hom}(V,W)$ come costruita nella Osservazione 6.3.3. È immediato verificare che $\Phi(f_A) = A$. Infatti, $\Phi(f_A)$ è la matrice che ha come i-esima colonna le m-upla ottenuta moltiplicando A per le coordinate dell'i-esimo vettore b_i della base $\mathcal{B}$, che sono le n-uple $e_i = (0, \dots, 1, \dots, 0)$. Ma il risultato di questa moltiplicazione è proprio la i-esima colonna di A. Quindi ogni matrice A è immagine di almeno una applicazione lineare, dunque la Φ è suriettiva.

3) Φ è iniettiva.
Poiché Φ è lineare, basta verificare che $\mathrm{Ker}\Phi = \{0\}$. Sia $f \neq 0$. (Qui 0 indica l'omomorfismo nullo). Allora $\exists b_i \in \mathcal{B}$ tale che $f(b_i) \neq 0_W$, perché

se le immagini dei vettori della base fossero tutti nulli, sarebbe $f = 0$ per il Lemma 6.1.8. Ma allora, essendo $f(b_i) \neq 0_W$, le coordinate di $f(b_i)$ non sono tutte nulle e quindi la matrice $\Phi(f)$ ha una colonna non nulla e perciò $\Phi(f) \neq 0$. (Qui 0 è la matrice nulla). Quindi f non sta in $\mathrm{Ker}\Phi$. □

Corollario 6.3.8 *Se V e W sono spazi vettoriale di dimensione n e m, allora* $\dim \mathrm{Hom}(V, W) = mn$.

Dim. Poiché $\mathrm{Hom}(V, W)$ è isomorfo a $\mathbf{M}_{m,n}(\mathbb{R})$, i due spazi hanno la stessa dimensione, che per $\mathbf{M}_{m,n}(\mathbb{R})$ sappiamo essere proprio mn. □

Un altro importante risultato che si inquadra in questo contesto, rigurda la composizione di applicazioni lineari, che sappiamo essere anch'essa lineare.

Proposizione 6.3.9 *Siano $f : V \to W$ e $g : W \to U$ due applicazioni lineari. Allora la funzione composta $g \circ f : V \to U$ è anch'essa lineare. Se inoltre vengono fissate delle basi su V, W e U, risulta: $\Phi(g \circ f) = \Phi(g)\Phi(f)$. Cioè la composizione di applicazioni lineari si traduce nel prodotto riga per colonna tra le matrici associate.*

Dim. La linearità di $f \circ g$ è semplice da verificare. Se $v_1, v_2 \in V$, $\lambda, \mu \in \mathbb{R}$,

$$\begin{aligned}(f \circ g)(\lambda v_1 + \mu v_2) &= g(f(\lambda v_1 + \mu v_2)) = g(\lambda f(v_1) + \mu f(v_2)) \\ &= \lambda g(f(v_1)) + \mu g(f(v_2)) = \lambda (fg)(v_1) + \mu (fg)(v_2)\,.\end{aligned}$$

Se X è il vettore colonna che esprime le coordinate di $v \in V$, $\Phi(f \circ g)X$ è il vettore colonna che esprime le coordinate di $(f \circ g)(v) = g(f(v)) \in U$. Le coordinate di $w = f(v) \in W$ sono date dal vettore colonna $Y = \Phi(g)X$, mentre quelle di $f(g(v))$ sono date da $\Phi(f)Y$. In definitiva, si ha:

$$\Phi(f \circ g)X = \Phi(f)Y = \Phi(f)\Phi(g)X \quad \forall X \in \mathbb{R}^n\,.$$

Quindi $(\Phi(f \circ g) - \Phi(f)\Phi(g))X = 0$ per ogni $X \in \mathbb{R}^n$. Ciò implica che: $(\Phi(f \circ g) - \Phi(f)\Phi(g)) = 0$ e infine $\Phi(f \circ g) = \Phi(f)\Phi(g)$. □

Osservazione 6.3.10 *Se* $\mathrm{id} : V \to V$ *è l'applicazione identità, cioè* $\mathrm{id}(v) = v$ *per ogni $v \in V$, e, cosa fondamentale, fissiamo la stessa base $\mathcal{B}$ sia quando V lo consideriamo come spazio di arrivo sia come spazio di partenza, allora: $\Phi(\mathrm{id}) = I_n$, dove I_n è la matrice identità di ordine n.*

Proposizione 6.3.11 *Sia $f : V \to W$ una applicazione lineare invertibile e sia $f^{-1} : W \to V$ la funzione inversa. La funzione f^{-1} è anch'essa lineare. Se inoltre vengono fissate delle basi su V e W, risulta: $\Phi(f^{-1}) = \Phi(f)^{-1}$. Cioè, la matrice associata alla inversa della applicazione lineare invertibile f è la matrice inversa della matrice associata alla f.*

Dim. Per prima cosa dimostriamo la linearità di $f^{-1} : W \to V$.
Siano $w_1, w_2 \in W$, $\lambda, \mu \in \mathbb{R}$, $v_1, v_2 \in V$ tali che $w_i = f(v_i)$; sia inoltre $v = f^{-1}(\lambda w_1 + \mu w_2)$ e $u = \lambda f^{-1}(w_1) + \mu f^{-1}(w_2)$.
Basta verificare che $u = v$. Infatti, si ha che:

$$f(u) = f(\lambda f^{-1}(w_1) + \mu f^{-1}(w_2)) = \lambda f(f^{-1}(w_1)) + \mu f(f^{-1}(w_2)) = \lambda w_1 + \mu w_2 \,.$$

Quindi, applicando f^{-1} ad entrambi i membri si ha:

$$u = f^{-1}(f(u)) = f^{-1}(\lambda w_1 + \mu w_2) = v \,.$$

Per la Proposizione 6.3.9, $\Phi(f \circ f^{-1}) = \Phi(f^{-1})\Phi(f)$ e, per l'Osservazione 6.3.10, $\Phi(f \circ f^{-1}) = \Phi(\mathrm{id}) = I$. Quindi si ha che $\Phi(f^{-1})\Phi(f) = I$, cioè $\Phi(f^{-1}) = \Phi(f)^{-1}$. □

Affrontiamo ora la spinosa questione della dipendenza di A_f dalla scelta delle basi in V e W. Supponiamo quindi di aver fissato su V una base $\mathcal{B}$ e su W una base $\mathcal{C}$. Sia $A = \Phi(f)$ la matrice associata alla applicazione lineare f in tali basi. Ricordiamo il significato di A.

Se le coordinate di $v \in V$, rispetto alla base $\mathcal{B}$, sono la n-upla

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

allora le coordinate di $f(v) \in W$, rispetto alla base $\mathcal{C}$, sono la m-upla

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

ottenuta ponendo $Y = AX$.

Sia $\mathcal{B}'$ un'altra base di V e sia X' la n-upla delle coordinate di v rispetto alla base $\mathcal{B}'$. Per la Proposizione 5.5.3, le n-uple X e X' sono legate dalla relazione $X = M_{\mathcal{B}}^{\mathcal{B}'} X'$, dove $M_{\mathcal{B}}^{\mathcal{B}'}$ è la matrice di passaggio dalle coordinate su V ripetto alla base $\mathcal{B}'$ nelle coordinate rispetto a $\mathcal{B}$. Allora

$$Y = AX = AM_{\mathcal{B}}^{\mathcal{B}'} X' .$$

Quindi *la matrice associata a f nelle basi $\mathcal{B}'$ di V e $\mathcal{C}$ di W è $AM_{\mathcal{B}}^{\mathcal{B}'}$.*

Analogamente, supponiamo di scegliere una base $\mathcal{C}'$ di W e sia $N_{\mathcal{C}'}^{\mathcal{C}}$ la matrice di passaggio dalle coordinate su W rispetto alla base $\mathcal{C}$ alle coordinate rispetto alla base $\mathcal{C}'$.

Se Y è la m-upla delle coordinate di $w \in W$ rispetto a $\mathcal{C}$, $Y' = N_{\mathcal{C}'}^{\mathcal{C}} Y$ è la m-upla delle coordinate dello stesso w nella base $\mathcal{C}'$. Quindi le coordinate di $f(v)$ nella base $\mathcal{C}'$ saranno:

$$Y' = N_{\mathcal{C}'}^{\mathcal{C}} Y = AX .$$

In definitiva abbiamo verificato il seguente enunciato:

Proposizione 6.3.12 *Sia $f : V \to W$ una applicazione lineare tra gli spazi vettoriali V e W sui quali è stata fissata una base, $\mathcal{B}$ su V e $\mathcal{C}$ su W. Sia A la matrice associata a f in tali basi. Se cambiamo basi, scegliendo $\mathcal{B}'$ su V e $\mathcal{C}'$ su W, la matrice A' associata alla stessa applicazione f rispetto a queste nuove basi è:*

$$A' = N_{\mathcal{C}'}^{\mathcal{C}} A M_{\mathcal{B}}^{\mathcal{B}'} , \tag{6.1}$$

dove $N_{\mathcal{C}'}^{\mathcal{C}}$ è la matrice di passaggio dalla base $\mathcal{C}$ a $\mathcal{C}'$ in W e $M_{\mathcal{B}}^{\mathcal{B}'}$ è la matrice di passaggio dalla base $\mathcal{B}'$ alla base $\mathcal{B}$ in V.

Possiamo, a questo punto, sospettare che, data una applicazione lineare $f : V \to W$, esistano delle basi privilegiate su V e W, rispetto alle quali la matrice associata alla f abbia forma particolarmente semplice.

Proposizione 6.3.13 *Sia $f : V \to W$ una applicazione lineare e sia $r =$ $\dim \mathrm{Im} f$. Allora esistono opportune basi di V e W rispetto alle quali la matrice associata alla f assume la seguente semplice forma:*

$$A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

dove I_r è la matrice identità di ordine r e gli 0 rappresentano blocchi di zeri di dimensione opportuna affinché A risulti di tipo $m \times n$, come richiesto.

Dim. Sia $\{b_1, \dots, b_{n-r}\}$ una base di Ker$f$ e sia $\{v_1, \dots, v_r\}$ un suo completamento ad una base $\mathcal{B} = \{v_1, \dots, v_r, b_1, \dots, b_{n-r}\}$ di V. Siano $w_1, \dots, w_r$ elementi di W tali che $w_i = f(v_i)$. Dalla dimostrazione del Teorema 6.2.9 possiamo affermare che $\{w_1, \dots, w_r\}$ sono una base di Imf ed in particolare, sono linearmente indipendenti. Completiamo ad una base $\mathcal{C}$ di W, $\mathcal{C} = \{w_1, \dots, w_r, u_1, \dots, u_{m-r}\}$. Scriviamo la matrice A associata alla f in queste basi.

Per $i = 1, \dots, r$, poiché $w_i = f(v_i)$, la i-esima colonna di A è il vettore e_i con il valore 1 all'i-esimo posto e zero altrimenti.

Le colonne di A dalla $r+1$-esima alla n-esima, essendo esse le coordinate di $f(b_i) = 0_W$, sono tutte nulle. Quindi A assume la forma predetta. □

Questo risultato, benché possa sembrare straordinario, non è molto utile nelle applicazioni perché il contesto in cui si lavora più spesso è il caso di endomorfismi $f : V \to V$, per i quali, scelta una base su V come spazio di partenza, il problema poi impone di scegliere la stessa base su V come spazio di arrivo, mentre è fondamentale nel risultato appena enunciato, poter liberamente cambiare sia la base di V che quella di W. Scegliere la base ottimale per scrivere la matrice di un endomorfismo è un problema ben più complesso che richiederà una trattazione a parte.

Poter lavorare in coordinate e scrivere le matrici associate alle applicazioni lineari permette di risolvere svariati problemi utilizzando calcoli espliciti.

Osservazione 6.3.14 *Sia $f : V \to W$ una applicazione lineare. Supponiamo di aver fissato una base $\mathcal{B}$ su V e una base $\mathcal{C}$ su W. Sia A la matrice associata a f rispetto ad esse. Allora:*

1. *$v \in$ Ker$f$ se e solo se le coordinate X di v rispetto alla base $\mathcal{B}$ sono soluzioni del sistema omogeneo $AX = 0$ di m equazioni in n incognite.*

2. *Nella stessa situazione, sia $w \in W$ e sia Y la m-upla delle sue coordinate. L'insieme $f^{-1}(w) = \{v \in V \mid f(v) = w\}$ è caratterizzato nel modo seguente: $v = f^{-1}(w)$ se e solo se le coordinate X di v sono soluzioni del sistema lineare $AX = Y$.*

3. *Se $A_{i_1}, \dots, A_{i_k}$ sono k colonne di A che concorrono nella formazione di un minore con determinante non nullo di ordine massimo, allora i k vettori di W che hanno come coordinate le m-uple $A_{i_1}, \dots, A_{i_k}$, sono una base di* Imf.

Infatti, sia $\tilde{A}$ *la matrice* $m \times k$ *formata dalle colonne* A_{i_j}. *Consideriamo le combinazioni lineari di queste colonne che danno il vettore nullo:* $x_1 A_{i_1} + \cdots + x_k A_{i_k} = 0$. *Le* k*-uple dei coefficienti sono le soluzioni del sistema omogeneo* $\tilde{A}X = 0$ *di* m *equazioni in* k *incognite, con* $k \leq m$. *Poiché* $\tilde{A}$ *ha rango massimo* k, *tale sistema ammette, per il teorema di Rouché-Capelli, solo la soluzione banale e pertanto le colonne sono linearmente indipendenti. Non solo. Ogni altra colonna* A_h *di* A *si ottiene come combinazione delle* A_{i_j}. *Infatti, la matrice* $(\tilde{A}, A_h)$ *è sempre una sottomatrice di* A, *perciò ha rango* k. *Sempre per Rouché-Capelli, il sistema* $\tilde{A}X = A_h$ *ha soluzioni e tali soluzioni sono i coefficienti della combinazione delle* A_{i_j} *che esprime la* A_h. *Per il Corollario 5.3.12 e l'Osservazione 6.2.7 si ha la tesi.*

4. *Il rango della matrice* A *associata alla applicazione lineare* f *è uguale alla dimensione dell'immagine di* f. *In formula:*

$$\mathrm{car} A = \dim \mathrm{Im} f \,. \tag{6.2}$$

Questo fatto segue subito dal punto precedente.

Data una matrice $A \in \mathbf{M}_{m,n}(\mathbb{R})$, essa determina un'applicazione lineare $f_A : \mathbb{R}^n \to \mathbb{R}^m$ definita da $f_A(X) = AX$. Per quanto appena detto, la caratteristica di A è allora esattamente uguale alla dimensione dell'immagine di f_A. Usando questo fatto è un facile esercizio dimostrare quanto segue.

Proposizione 6.3.15 *Se* A *e* B *sono due matrici per cui si possa fare il prodotto riga per colonna* AB, *allora:*

$$\mathrm{car}(AB) \leq \min\{\mathrm{car} A, \mathrm{car} B\} \,. \tag{6.3}$$

Se B *è invertibile, allora* $\mathrm{car}(AB) = \mathrm{car} A$. *Analogamente, se* A *è invertibile, allora* $\mathrm{car}(AB) = \mathrm{car} B$.

6.4 Esercizi

Applicazioni lineari

1) Date le matrici

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 4 & 3 & 1 \\ 5 & 2 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 & 1 & 2 \\ 1 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 \\ 2 & 1 & 1 & 2 \end{pmatrix},$$

si calcolino il nucleo e l'immagine delle applicazioni lineari $L_A : \mathbb{R}^3 \to \mathbb{R}^3$ e $L_B : \mathbb{R}^4 \to \mathbb{R}^4$ associate ad A e B (rispetto alle basi canoniche di $\mathbb{R}^3$ e $\mathbb{R}^4$, rispettivamente).

2) Costruire, se possibile, un'applicazione lineare $F : \mathbb{R}^3 \to \mathbb{R}^3$ la cui immagine sia generata dai vettori

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}.$$

3) Costruire, se possibile, un'applicazione lineare $F : \mathbb{R}^4 \to \mathbb{R}^2$ il cui nucleo sia generato dai vettori

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

4) Dati

$$u_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 1 \\ 7 \\ 9 \end{pmatrix}, \quad v_2 = \begin{pmatrix} \pi \\ 3 \\ 5 \end{pmatrix},$$

costruire, se possibile, un'applicazione lineare $F : \mathbb{R}^4 \to \mathbb{R}^3$ il cui nucleo sia generato da u_1 e u_2 e la cui immagine sia generata da v_1 e v_2.

5) Data la matrice

$$A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 2 & 2 \\ 3 & 1 & k \end{pmatrix}$$

con k parametro reale, sia $F = L_A$ l'applicazione lineare associata ad A rispetto alla base canonica di $\mathbb{R}^3$. Dire per quali valori di k esiste un'applicazione lineare $G : \mathbb{R}^3 \to \mathbb{R}$ suriettiva tale che $G \circ F = 0$. Nei casi in cui è possibile costruire una tale G, lo si faccia.

6) Trovare, se possibile, un'applicazione $F : \mathbb{R}^3 \to \mathbb{R}^3$ lineare tale che:

$$F\begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \quad F\begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad F\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix}.$$

7) Determinare per quali valori del parametro reale k esiste un'applicazione $F : \mathbb{R}^3 \to \mathbb{R}^3$ lineare tale che:

$$F\begin{pmatrix} 2 \\ 3 \\ 1+k \end{pmatrix} = \begin{pmatrix} 3 \\ k \\ 0 \end{pmatrix}, \quad F\begin{pmatrix} 1 \\ 0 \\ k \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}, \quad F\begin{pmatrix} 6 \\ k+9 \\ 3 \end{pmatrix} = \begin{pmatrix} 9 \\ 3k \\ k \end{pmatrix}.$$

8) Sia V uno spazio vettoriale di dimensione n e F un endomorfismo nilpotente (cioè $F : V \to V$ è un'applicazione lineare per cui esiste un m tale che $F^m = 0$). Dimostrare che, se $F^{n-1} \neq 0$, allora esiste $v \in V$ tale che $v, F(v), \dots, F^{n-1}(v)$ è una base di V.

9) Considerato lo spazio vettoriale $\mathbb{R}_2[x]$ dei polinomi a coefficienti reali di grado al più due, sia $F : \mathbb{R}_2[x] \to \mathbb{R}_2[x]$ l'applicazione tale che:

$$F(p)(x) = p(x-1) - xp'(x).$$

Stabilire se F è lineare e, in caso affermativo, trovarne il nucleo e l'immagine.

10) Fissato $\lambda \in \mathbb{R}$, sia $F_\lambda : \mathbb{R}_2[x] \to \mathbb{R}^3$ l'applicazione definita da

$$F_\lambda(p) = {}^t(p(0), p(1), p(\lambda)).$$

Mostrare che l'applicazione F è lineare e determinare per quali valori del parametro λ essa è biunivoca.

11) Costruire, se possibile, un'applicazione lineare $F : \mathbb{R}^8 \to M_3(\mathbb{R})$ tale che $\mathrm{Im}(F) \supseteq \{A \in M_3(\mathbb{R}) \mid \mathrm{rk}(A) = 2\}$.

Matrici associate ad applicazioni lineari

12) Sia $F : \mathbb{R}^2 \to \mathbb{R}^3$ l'applicazione definita da

$$F\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x-y \\ -y \end{pmatrix} .$$

Si scrivano:
a) la matrice $M(F)$ associata a F rispetto alle basi canoniche di $\mathbb{R}^2$ e $\mathbb{R}^3$;
b) la matrice $M^{\mathcal{B}}_{\mathcal{C}}(F)$ associata a F rispetto alle seguenti basi $\mathcal{B}$ e $\mathcal{C}$

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}, \quad \mathcal{C} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\} .$$

13) Sia data la base di $\mathbb{R}^3$

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\} .$$

Indicata con $\mathcal{C}an = \{e_1, e_2, e_3\}$ la base canonica di $\mathbb{R}^3$, si calcolino la matrici $M^{\mathcal{C}an}_{\mathcal{C}an}(\mathrm{id}_{\mathbb{R}^3})$, $M^{\mathcal{B}}_{\mathcal{B}}(\mathrm{id}_{\mathbb{R}^3})$, $M^{\mathcal{B}}_{\mathcal{C}an}(\mathrm{id}_{\mathbb{R}^3})$, $M^{\mathcal{C}an}_{\mathcal{B}}(\mathrm{id}_{\mathbb{R}^3})$, associate all'applicazione identità $\mathrm{id}_{\mathbb{R}^3}$ di $\mathbb{R}^3$.

14) Sia $F : \mathbb{R}^3 \to \mathbb{R}^3$ l'applicazione definita da

$$F\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x-y-2z \\ x+y+z \\ 3x-z \end{pmatrix} .$$

Si scrivano:
a) la matrice $M(F) = M^{\mathcal{C}an}_{\mathcal{C}an}(F)$ associata a F rispetto alla base canonica $\mathcal{C}an$ di R^3;
b) la matrice $M^{\mathcal{B}}_{\mathcal{B}}(F)$ associata a F rispetto alla seguente base

$$\mathcal{B} = \left\{ \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \right\} .$$

Capitolo 7

DIAGONALIZZAZIONE DI MATRICI E FORMA CANONICA DI JORDAN

7.1 Similitudine

In tutto questo capitolo, salvo esplicita avvertenza del contrario, le matrici saranno quadrate.

Definizione 7.1.1 *Due matrici* A *e* B *si dicono* simili *se esiste una matrice invertibile* M *tale che:*

$$B = M^{-1}AM\,.$$

Proposizione 7.1.2 *La relazione di similitudine sull'insieme* $\mathbf{M}_n(\mathbb{R})$ *delle matrici quadrate reali di ordine* n *è una relazione di equivalenza.*

Dim. 1) Proprietà riflessiva. Basta prendere $M = I_n$.
2) Proprietà simmetrica. Supponiamo che A sia simile a B. Abbiamo $B = M^{-1}AM$. Moltiplicando a sinistra per M e a destra per M^{-1}, si ha che $A = MBM^{-1}$. Posto $N = M^{-1}$, si ha che $A = N^{-1}BN$. Ossia B è simile ad A.
3) Proprietà transitiva. Supponiamo che $B = M^{-1}AM$ e che $C = N^{-1}BN$. Allora, sostituendo, $C = N^{-1}(M^{-1}AM)N = (MN)^{-1}AMN$. Posto $MN = G$, $C = G^{-1}AG$, quindi anche A e C sono simili. □

Osservazione 7.1.3 *La matrice nulla* 0 *e l'identità* I_n *sono simili solo a se stesse. Infatti:* $B = M^{-1}0M = 0$; $B = M^{-1}I_nM = M^{-1}M = I_n$. *Similmente si dimostra che le matrici del tipo* λI_n, *con* λ *un fissato numero reale, sono simili solo a se stesse.*

Proposizione 7.1.4 *Se due matrici* A *e* B *sono simili, allora hanno lo stesso determinante.*

Dim. Supponiamo $B = M^{-1}AM$, con M invertibile. Ricordiamo che $\det(M^{-1}) = (\det M)^{-1}$. Allora, sempre per la formula di Binet, abbiamo

$$\det B = \det(M^{-1}AM) = \det(M^{-1}) \det A \det M = \det A\,.$$

□

Avere lo stesso determinante è condizione necessaria per essere simili ma in generale non sufficiente. In altre parole due matrici simili hanno lo stesso determinante, ma due matrici (di ordine maggiore di uno) possono avere lo stesso determinante e non essere simili.

Esempio 7.1.5 *Siano*

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \; e \; B = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

Si ha $\det A = \det B = 4$. *Ma esse non sono simili. Supponiamo che lo siano; allora esisterebbe* $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *tale che* $B = M^{-1}AM$, *che è equivalente a dire che* $MB = AM$ *e* $\det M \neq 0$. *Abbiamo:*

$$\begin{aligned} MB &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2a & a+2b \\ 2c & c+2d \end{pmatrix}, \\ AM &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 2a & 2b \\ 2c & 2d \end{pmatrix}. \end{aligned}$$

Quindi $AM = MB$ *implica che*

$$\begin{pmatrix} 2a & 2b \\ 2c & 2d \end{pmatrix} = \begin{pmatrix} 2a & a+2b \\ 2c & c+2d \end{pmatrix}.$$

Due matrici sono uguali se e solo se hanno tutti gli elementi uguali. L'uguaglianza tra le due matrici si traduce nel sistema:

$$\begin{cases} 2a = 2a\,, \\ 2c = 2c\,, \\ a + 2b = 2b\,, \\ c + 2d = 2d\,, \end{cases}$$

che ha come soluzioni $a = 0, c = 0$. *Ma allora* M *deve avere la prima colonna nulla e dunque non può essere invertibile. Quindi* A *e* B *non sono simili.*

Questa è la verifica diretta da usarsi in generale. In questo caso specifico, bastava osservare che $A = 2I_2$ *che non può essere simile ad una matrice diversa da se stessa.*

Abbiamo visto che, fissate delle basi, una applicazione lineare tra spazi vettoriali è rappresentata da una matrice.

Sia $f : V \to V$ un endomorfismo, cioè una applicazione lineare di uno spazio V in sé, e sia A la matrice associata ad f quando sia stata fissata una base $\mathcal{B}$ di V, inteso sia come spazio di partenza sia come spazio di arrivo. La matrice A è quadrata di ordine uguale alla dimensione di V.

La matrice A', associata ad f in un'altra base $\mathcal{B}'$ di V, si ricava applicando la Proposizione 6.3.12 al caso in cui: $W = V, \mathcal{B} = \mathcal{C}, \mathcal{B}' = \mathcal{C}'$. Si ha:

$$A' = M^{\mathcal{B}}_{\mathcal{B}'} A M^{\mathcal{B}'}_{\mathcal{B}}\,. \tag{7.1}$$

dove $M^{\mathcal{B}'}_{\mathcal{B}}$ è la matrice di passaggio dalla base $\mathcal{B}'$ alla base $\mathcal{B}$ su V. Ma ricordando che $M^{\mathcal{B}}_{\mathcal{B}'} = (M^{\mathcal{B}'}_{\mathcal{B}})^{-1}$, si trova che A e A' sono simili. Dunque:

Osservazione 7.1.6 *Due matrici quadrate sono simili se e solo se rappresentano, in basi diverse, lo stesso endomorfismo.*

Con ciò è possibile definire il *determinante di un endomorfismo.*

Se f è un endomorfismo, $\det(f) = \det A$, dove A è una matrice che rappresenta f in una qualsiasi base. Questa definizione ha senso poiché ogni altra matrice che rappresenti f, essendo simile ad A, ha lo stesso determinante.

Osservazione 7.1.7 *Sia* A *una matrice che rappresenta l'endomorfismo* f *in una qualche base. Ricordiamo che* $\mathrm{car}A = \dim \mathrm{Im} f$. *Se* A' *è una matrice simile ad* A, *essa rappresenta* f *in un'altra base, dunque* $\mathrm{car}A' = \dim \mathrm{Im} f$. *Si può quindi dedurre che* due matrici simili hanno la stessa caratteristica.

Per descrivere un endomorfismo sarà opportuno scegliere la base nella quale la matrice associata avrà forma particolarmente semplice. Il problema che ci accingiamo a risolvere è il seguente:

data una matrice A, determinare nella classe di equivalenza per similitudine di A la matrice che sia, in un certo senso, la più semplice di tutte, ad esempio che abbia il maggior numero di elementi uguali a zero.

7.2 Autovettori, autovalori e polinomio caratteristico

Definizione 7.2.1 *Sia A una matrice (reale) quadrata di ordine n. Il vettore colonna* $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, *$X \neq 0$, è un* autovettore *di A, se esiste uno scalare λ tale che:*

$$AX = \lambda X\,. \tag{7.2}$$

Definizione 7.2.2 *Il numero (reale) λ è un* autovalore *della matrice A se esiste un vettore colonna X,* non nullo, *tale che $AX = \lambda X$.*

Esempio 7.2.3

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}\,, \quad X = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}\,,$$
$$AX = \begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}\,.$$

Si dice che X è autovettore di A relativo all'autovalore 2.

Proposizione 7.2.4 *Data A matrice quadrata, dato $\lambda \in \mathbb{R}$, esistono autovettori relativi a λ se e solo se:*

$$\det(A - \lambda I_n) = 0\,.$$

Dim. Per definizione, X è autovettore relativo a λ se $X \neq 0$ e $AX = \lambda X$. Ciò equivale a dire che $(A - \lambda I_n)X = 0$ (qui 0 rappresenta la colonna fatta tutta di zeri). Una siffatta n-upla X è dunque una soluzione del sistema scritto in forma matriciale $(A - \lambda I_n)X = 0$ di cui $(A - \lambda I_n)$ rappresenta la matrice dei coefficienti. Per il teorema di Cramer, questo sistema ammette soluzioni non nulle se e solo se $\det(A - \lambda I_n) = 0$. □

Sia ora t una indeterminata.

Definizione 7.2.5 $\det(A - tI_n)$ *è un polinomio di grado n nella indeterminata t e viene detto* polinomio caratteristico *di A. Viene indicato con $p_A(t)$.*

Che $p_A(t) = \det(A - tI_n)$ sia un polinomio lo si deduce dal fatto che la matrice $A - tI_n$ ha i termini sulla diagonale principale che sono polinomi di primo grado in t. Siccome somme e prodotti tra polinomi danno sempre polinomi e siccome il determinante di una matrice si calcola facendo prodotti tra elementi della matrice stessa e sommando i vari risultati, tale determinante sarà un polinomio. Inoltre, si dimostra (per induzione) che il grado di $p_A(t)$ è esattamente uguale a n, se n è l'ordine di A.

Proposizione 7.2.6 *Se A e B sono simili, esse hanno lo stesso polinomio caratteristico.*

Dim. Se $B = M^{-1}AM$, allora

$$B - tI = M^{-1}AM - tI = M^{-1}AM - tM^{-1}IM = M^{-1}(A - tI)M\,.$$

Dunque:

$$\begin{aligned} p_B(t) &= \det(B - tI) = \det(M^{-1}(A - tI)M) \\ &= \det(M^{-1})\det(A - tI)\det M = \det(A - tI) = p_A(t)\,. \end{aligned}$$

□

Osservazione 7.2.7 *Non vale il viceversa: due matrici aventi lo stesso polinomio caratteristico possono non essere simili.*
Consideriamo A e B dell'Esempio 7.1.5. Sappiamo che A e B non sono simili, ma:

$$\begin{aligned} p_A(t) &= \det\begin{pmatrix} 2-t & 0 \\ 0 & 2-t \end{pmatrix} = (2-t)^2\,, \\ p_B(t) &= \det\begin{pmatrix} 2-t & 1 \\ 0 & 2-t \end{pmatrix} = (2-t)^2\,. \end{aligned}$$

Osservazione 7.2.8 *Le radici del polinomio caratteristico della matrice A sono tutti e soli gli autovalori di A. Infatti se λ è una radice del polinomio caratteristico, allora sostituendo λ al posto di t si ha zero, cioè* $\det(A-\lambda I) = 0$. *Ma allora, per Rouché-Capelli, il sistema scritto in forma matriciale: $(A - \lambda I)X = 0$ ammette soluzioni non nulle, che sono, per definizione, gli autovettori di A. Ripercorrendo a ritroso il ragionamento si dimostra l'altra implicazione.*

In base a questa osservazione, una matrice di ordine n ha al massimo n autovalori.

Definizione 7.2.9 *Diremo che l'autovalore λ ha* molteplicità algebrica *k se λ è radice di molteplicità k (nel senso della Definizione 2.3.9) del polinomio caratteristico di A. La molteplicità algebrica dell'autovalore λ si indica con $m_a(\lambda)$.*

Osservazione 7.2.10 *Se tutti gli autovalori di A sono reali, il polinomio caratteristico, in virtù del teorema di Ruffini, si può scomporre in fattori del tipo $(t - \lambda)^{m_a(\lambda)}$. Dunque la somma delle molteplicità algebriche di tutti gli autovalori distinti di A è uguale a n.*

La matrice A dell'Esempio 7.1.5 ammette 2 come autovalore. Siccome $p_A(t) = (2-t)^2$, si ha $m_a(2) = 2$.

Definizione 7.2.11 *Diremo che l'autovalore λ ha* molteplicità geometrica *k se $k = n - car(A - \lambda I)$, dove n è l'ordine di A. La molteplicità geometrica dell'autovalore λ si indica con $m_g(\lambda)$.*

Esempio 7.2.12 *1)*

$$A = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \quad A - tI = \begin{pmatrix} 2-t & 0 & -1 \\ 0 & 2-t & 2 \\ 0 & 0 & 1-t \end{pmatrix}.$$

$$p_A(t) = \det(A - tI) = (2-t)^2(1-t).$$

La matrice A ha due autovalori distinti 2 e 1 con $m_a(2) = 2$ e $m_a(1) = 1$. Essendo

$$A - 2I = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & -1 \end{pmatrix},$$

chiaramente car$(A-2I)=1$ *e dunque* $m_g(2)=3-1=2$.
2)

$$A=\begin{pmatrix}-3 & 1\\ 0 & -3\end{pmatrix}, \quad A-tI=\begin{pmatrix}-3-t & 1\\ 0 & -3-t\end{pmatrix}.$$

$$p_A(t)=(-3-t)^2, \quad m_a(-3)=2.$$

Calcoliam la molteplicità geometrica di -3. $\mathrm{car}(A+3I)=\mathrm{car}\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}=1$. *Allora* $m_g(-3)=2-1=1$.

In generale si può dimostrare che $1\leq m_g(\lambda)\leq m_a(\lambda)$. Cosa che faremo in seguito.

Definizione 7.2.13 *Un autovalore* λ *è detto* regolare *se accade che:*

$$m_a(\lambda)=m_g(\lambda).$$

Teorema 7.2.14 *L'insieme di tutti gli autovettori relativi ad un dato autovalore* λ *di una matrice* A*, unito a* $\{0\}$*, è un sottospazio vettoriale di* $\mathbb{R}^n$*. Tale sottospazio si chiama* autospazio *relativo all'autovalore* λ*, si indica con* $\mathrm{Aut}(\lambda)$ *e ha dimensione* $m_g(\lambda)$.

Dim. Per definizione, $0\in\mathrm{Aut}(\lambda)$. Possiamo dire che $X\in\mathrm{Aut}(\lambda)$ se $AX=\lambda X$ senza che sia richiesto $X\neq 0$.

1. Se $X\in\mathrm{Aut}(\lambda)$ e $t\in\mathbb{R}$, $tX\in\mathrm{Aut}(\lambda)$.
   Infatti $A(tX)=tAX=t\lambda X=\lambda(tX)$.

2. Se $X,Y\in\mathrm{Aut}(\lambda)$, allora $X+Y\in\mathrm{Aut}(\lambda)$.
   Infatti $A(X+Y)=AX+AY=\lambda X+\lambda Y=\lambda(X+Y)$.

Sinteticamente, $\mathrm{Aut}(\lambda)=\mathrm{Ker}(A-\lambda I)$ (più esattamente, Ker$f$, dove f è l'endomorfismo di $\mathbb{R}^n$ associato alla matrice $A-\lambda I$ nella base canonica). Quindi, essendo il nucleo di un endomorfismo, è un sottospazio lineare.

$$\dim\mathrm{Aut}(\lambda)=\dim\mathbb{R}^n-dim\mathrm{Im}(A-\lambda I)=n-\mathrm{car}(A-\lambda I)=m_g(\lambda).$$

□

Proposizione 7.2.15 *Se λ e μ sono autovalori di A distinti, allora:*

$$\mathrm{Aut}(\lambda) \cap \mathrm{Aut}(\mu) = \{0\}\,.$$

Dim. Sia $X \neq 0$, $X \in \mathrm{Aut}(\lambda) \cap \mathrm{Aut}(\mu)$. Poiché $X \in \mathrm{Aut}(\lambda)$, $AX = \lambda X$. Analogamente $AX = \mu X$. Ma allora $AX - AX = \lambda X - \mu X = (\lambda - \mu)X = 0$. Siccome $\lambda \neq \mu$, $\lambda - \mu \neq 0$ e avendo supposto $X \neq 0$, si ha un assurdo. □

Corollario 7.2.16 *Autovettori di autovalori distinti sono sempre linearmente indipendenti.*

7.3 Diagonalizzazione di una matrice

Anche se una matrice è ad elementi reali e quindi il suo polinomio caratteristico è a coefficienti reali, potrebbe comunque avere autovalori complessi.

Definizione 7.3.1 *Una matrice quadrata A è detta matrice* diagonale *se ha tutti gli elementi al di fuori della sua diagonale principale uguali a zero. Ossia, $a_{ij} = 0$ se $i \neq j$.*

Una matrice diagonale è quindi individuata dando nell'ordine gli elementi che si trovano sulla sua diagonale principale. Indicheremo con $\mathrm{diag}(d_1, \ldots, d_n)$ la matrice tale che $a_{ii} = d_i$ e $a_{ij} = 0$ se $i \neq j$. Ad esempio

$$\mathrm{diag}(1, 0, -2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}\,.$$

Definizione 7.3.2 *Una matrice A è detta matrice* diagonalizzabile *se A è simile ad una matrice diagonale.*

Un endomorfismo f di uno spazio vettoriale V si dice diagonalizzabile *se esiste una base $\mathcal{B}$ di V tale che la matrice associata a f rispetto a $\mathcal{B}$ è diagonale. Questo accade se e solo se $\mathcal{B}$ è una base formata da autovettori di f.*

Chiaramente una matrice diagonale è la più semplice matrice con cui si possa sperare di lavorare. Purtroppo, come sarà presto evidente, non tutte le matrici sono diagonalizzabili. Cercheremo allora di determinare quali siano le condizioni per la diagonalizzabilità di una matrice.

Osservazione 7.3.3 *Una matrice diagonale ha tutti gli autovalori reali e regolari.*

Infatti sia $D = \mathrm{diag}(d_1, \ldots, d_n)$. *Essendo* $D - tI = \mathrm{diag}(d_1 - t, \ldots, d_n - t)$, *si ha:*

$$p_D(t) = \det(D - tI_n) = (d_1 - t)(d_2 - t) \cdots (d_n - t)\,.$$

Quindi $p_D(t)$ *ha* n *radici reali, proprio i valori della diagonale principale di* D, *eventualmente con molteplicità se i* d_i *non sono tutti distinti.*

La molteplicità algebrica di un autovalore λ *di* D *è in questo caso il numero di volte che* λ *compare sulla diagonale. Allora la matrice* $D - \lambda I$ *è anch'essa diagonale e ha sulla diagonale principale tanti zeri quante sono le volte che* λ *compariva sulla diagonale di* D, *cioè* $m_a(\lambda)$ *volte.*

Se cancello da $D - \lambda I$ *le righe e le colonne tutte nulle, che sono in numero uguale a* $m_a(\lambda)$, *ciò che resta è un minore* D' *che è diagonale e con gli elementi sulla sua diagonale tutti non nulli. Pertanto il suo determinante è diverso da zero. Quindi* $car(D - \lambda I) = n - m_a(\lambda)$.

Concludiamo dicendo che, per ogni autovalore λ, *si ha:*

$$m_g(\lambda) = n - \mathrm{car}(D - \lambda I) = m_a(\lambda)\,,$$

quindi λ *è regolare.*

A questo punto possiamo provare il risultato fondamentale.

Teorema 7.3.4 *Una matrice* A *è diagonalizzabile se e solo se ha tutti gli autovalori reali e regolari.*

Dim. Se A è simile a una matrice diagonale D, A ha lo stesso determinante di D, la stessa caratteristica, lo stesso polinomio caratteristico e quindi gli stessi autovalori con la stessa molteplicità algebrica e anche geometrica (perché anche $A - \lambda I$ è simile a $D - \lambda I$ e quindi hanno la stessa caratteristica). Quindi, poiché D ha autovalori tutti reali e regolari, anche quelli di A lo sono.

Viceversa, sia A una matrice i cui autovalori sono tutti reali e regolari.

Siano $\lambda_1, \ldots, \lambda_h$ gli autovalori distinti di A. Consideriamo $\mathrm{Aut}(\lambda_1) \oplus \cdots \oplus \mathrm{Aut}(\lambda_h)$. Possiamo scriverli in somma diretta perché la loro intersezione a due a due si riduce al solo $\{0\}$ (cfr. Proposizione 7.2.15). Poiché sono tutti reali, $m_a(\lambda_1) + \cdots + m_a(\lambda_h) = n$. Poiché sono tutti regolari, anche $m_g(\lambda_1) + \cdots + m_g(\lambda_h) = n$. Per la relazione di Grassmann:

$$\begin{aligned}\dim(\mathrm{Aut}(\lambda_1) \oplus \cdots \oplus \mathrm{Aut}(\lambda_h)) &= \dim \mathrm{Aut}(\lambda_1) + \cdots + \dim \mathrm{Aut}(\lambda_h) \\ &= m_g(\lambda_1) + \cdots + m_g(\lambda_h) = n\,.\end{aligned}$$

Quindi

$$\mathrm{Aut}(\lambda_1) \oplus \cdots \oplus \mathrm{Aut}(\lambda_h) = \mathbb{R}^n .$$

Sia $\mathcal{B}$ una base di $\mathbb{R}^n$ ottenuta come unione di basi dei singoli $\mathrm{Aut}(\lambda_i)$. La base $\mathcal{B}$ è una fatta di autovalori di A. Sia M la matrice che ha come colonne i vettori di $\mathcal{B}$. Abbiamo che:
1) $\det M \neq 0$, poiché le colonne sono linearmente indipendenti;
2) $M^{-1}AM$ è diagonale.
Sia infatti d_{ij} l'elemento di posto (i,j) di $M^{-1}AM$. Esso si ottiene moltiplicando la i-esima riga di M^{-1} con la j-esima colonna di AM. La j-esima colonna di AM si ottiene moltiplicando A per la j-esima colonna di M. Sia X_j la j-esima colonna di M. Per come è stata costruita M, X_j è un autovettore di A e pertanto $AX_j = \lambda_k X_j$, per un certo autovalore λ_k. Quindi la j-esima colonna di AM è λ_k volte la j-esima colonna di M. Quindi, moltiplicare la i-esima riga di M^{-1} con la j-esima colonna di AM equivale a moltiplicare la i-esima riga di M^{-1} con la j-esima colonna di M e poi moltiplicare il risultato per il corrispondente autovalore. In ogni caso, se $i \neq j$, se moltiplico la i-esima riga di M^{-1} con la j-esima colonna di M ottengo, per la definizione di matrice inversa, zero.

Possiamo quindi concludere che, siccome l'elemento di posto (i,j) di $M^{-1}AM$ è zero, quindi $M^{-1}AM$ è una matrice diagonale. □

Il procedimento della dimostrazione precedente ha come immediate conseguenze i seguenti fatti.

Osservazione 7.3.5 *Una matrice A è diagonalizzabile se e solo se esiste una base di $\mathbb{R}^n$ composta da autovettori di A. In tal caso A è simile ad una matrice diagonale che ha sulla diagonale gli autovalori di A con molteplicità uguale a quella geometrica (o algebrica, visto che in questo caso è la stessa).*

Per costruire esplicitamente la matrice di passaggio che definisce la similitudine tra A e una matrice diagonale D basta prendere la matrice M che ha come colonne una base (ordinata) di autovettori di A. L'odine in cui compaiono gli autovalori di A sulla diagonale della matrice $D = M^{-1}AM$ dipende esattamente dall'ordine scelto per gli autovettori che formano la base $\mathcal{B}$.

Per fare ciò bisogna seguire il seguente schema.

Schema 3 *Sia data la matrice A.*
1) Calcolare il polinomio caratteristico di A.
2) Determinare gli autovalori di A (unico passaggio difficile).

3) Verificare che tutti gli autovalori siano regolari e che quindi A sia effettivamente diagonalizzabile.
4) Per ogni autovalore λ_k di A determinare $m_g(\lambda_k)$ autovettori indipendenti.
5) Scrivere la matrice M che ha come colonne tutti i vettori ottenuti iterando il punto 4) per tutti gli autovalori di A.
6) Verificare (facoltativo) che $M^{-1}AM$ è una matrice diagonale.

Spieghiamo più in dettaglio come operare effettivamente al punto 4). Sappiamo che gli autovettori relativi all'autovalore λ_k sono le soluzioni del sistema, scritto in forma matriciale:

$$(A - \lambda_k I)X = 0\,.$$

Applicando il metodo visto nello Schema 1, vediamo che l'insieme delle soluzioni dipende da $n - car(A - \lambda_k I) = m_g(\lambda_k)$ parametri. Supponiamo che $m_g(\lambda_k) = p$ e siano $t_1, \dots, t_p$ questi parametri. Chiamiamo X_i la soluzione che si ottiene ponendo $t_i = 1$ e $t_j = 0$ per $i \neq j$. Si può dimostrare che i p vettori X_i così ottenuti sono linearmente indipendenti, fatto che viene lasciato per esercizio.

Esempio 7.3.6 *Consideriamo*

$$A = \begin{pmatrix} -1 & 0 & 0 & 2 \\ -3 & 2 & 0 & 3 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Vogliamo stabilire se A è diagonalizzabile e, in caso affermativo, determinare una matrice invertibile M tale che $M^{-1}AM$ sia una matrice diagonale.

Seguiamo lo Schema 3.
1)

$$\begin{aligned} p_A(t) &= \det(A - tI) = \det \begin{pmatrix} -1-t & 0 & 0 & 2 \\ -3 & 2-t & 0 & 3 \\ 0 & 0 & -1-t & 0 \\ 0 & 0 & 0 & 1-t \end{pmatrix} \\ &= (1-t)(2-t)(-1-t)^2\,. \end{aligned}$$

2) Gli autovalori di A sono $\{1, -1, 2\}$.
3) $m_a(1) = m_a(2) = 1$. Poiché $1 \le m_g(\lambda) \le m_a(\lambda) = 1$, è chiaro che sia

l'autovalore 1 *che l'autovalore* 2 *sono regolari. Poiché* $m_a(-1) = 2$*, occorre invece determinare* $m_g(-1)$*. Essendo*

$$\mathrm{car}(A+I) = \mathrm{car}\begin{pmatrix} 0 & 0 & 0 & 2 \\ -3 & 3 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = 2\,,$$

abbiamo $m_g(-1) = 4 - 2 = 2$*. Quindi anche* -1 *è regolare e* A *è diagonalizzabile.*
4) Costruiamo ora una base di $\mathbb{R}^4$ *fatta da autovettori di* A*.*

Autovalore 1. *Risolviamo il sistema* $(A - I)X = 0$*:*

$$\begin{pmatrix} -2 & 0 & 0 & 2 \\ -3 & 1 & 0 & 3 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Un minore di ordine 3 *con determinante non nullo è* N *ottenuto cancellando la quarta riga e la quarta colonna,* $N = \begin{pmatrix} -2 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$.

La generica soluzione si ottiene risolvendo con Cramer il nuovo sistema ridotto:

$$\begin{pmatrix} -2 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -2w \\ -3w \\ 0 \end{pmatrix}.$$

La quarta incognita w *diventa parametro ponendo* $w = t$*. La generica soluzione è* $\begin{pmatrix} t \\ 0 \\ 0 \\ t \end{pmatrix}$*, da cui, prendendo* $t = 1$*,* $X_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$.

Autovalore 2. *Risolviamo il sistema* $(A - 2I)X = 0$*.*

$$\begin{pmatrix} -3 & 0 & 0 & 2 \\ -3 & 0 & 0 & 3 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Un minore di ordine 3 *con determinante non nullo è* N' *ottenuto cancellando la prima riga e la seconda colonna.* $N' = \begin{pmatrix} -3 & 0 & 3 \\ 0 & -3 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

La generica soluzione si ottiene risolvendo con Cramer il nuovo sistema ridotto:

$$\begin{pmatrix} -3 & 0 & 3 \\ 0 & -3 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

La seconda incognita y *diventa parametro ponendo* $y = t$. *La generica soluzione è* $\begin{pmatrix} 0 \\ t \\ 0 \\ 0 \end{pmatrix}$, *da cui, prendendo* $t = 1$, $X_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

Autovalore -1. *Risolviamo il sistema* $(A + I)X = 0$.

$$\begin{pmatrix} 0 & 0 & 0 & 2 \\ -3 & 3 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} .$$

Il rango di $A + I$ *è uguale a* $4 - m_g(-1) = 2$.

Un minore di ordine 2 *con determinante non nullo è* N'' *ottenuto cancellando prima e terza riga e seconda e terza colonna.* $N'' = \begin{pmatrix} -3 & 3 \\ 0 & 2 \end{pmatrix}$.

La generica soluzione si ottiene risolvendo con Cramer il nuovo sistema ridotto:

$$N'' = \begin{pmatrix} -3 & 3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} = \begin{pmatrix} -3y \\ 0 \end{pmatrix} .$$

La seconda e la terza incognita y *e* z *diventano parametri ponendo* $y = t_1$ *e* $z = t_2$. *La generica soluzione è*

$$\begin{pmatrix} t_1 \\ t_1 \\ t_2 \\ 0 \end{pmatrix} .$$

Questa volta abbiamo due autovettori indipendenti.

Prendendo $t_1 = 1, t_2 = 0$, *abbiamo* $X_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$.

Prendendo $t_1 = 0, t_2 = 1$, *abbiamo* $X_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$.

I vettori X_1, X_2, X_3, X_4 *così costruiti danno una base di* $\mathbb{R}^4$ *fatta di autovettori.*

5) $M = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$. *La i-esima colonna di* M *è data dal vettore* X_i.

Il punto 6) viene lasciato per esercizio.

Osserviamo che, per come abbiamo costruito M, *se* $D = M^{-1}AM$, $D = \mathrm{diag}(d_1, \dots, d_n)$, *si ha che* d_i *è l'autovalore relativo all'i-esimo autovettore* X_i *della base scelta.*

Nel nostro esempio, $M^{-1}AM = \mathrm{diag}(1, 2, -1, -1)$.

7.4 Forma canonica di Jordan

Abbiamo visto che due matrici sono simili se e solo se sono associate ad uno stesso endomorfismo, ovviamente in basi diverse.

La nostra strategia sarà la seguente: considereremo un endomorfismo F di V come rappresentante della classe di equivalenza delle matrici simili che rappresentano F nelle varie basi e studiandone le proprietà, individueremo una base di V in cui F assuma una forma che sia la più semplice.

Poiché matrici simili hanno, tra le altre cose, lo stesso polinomio caratteristico, ha dunque senso definire gli autovalori di un endomorfismo F come gli autovalori di una delle matrici associate a F.

Si può dimostrare che questa definizione coincide con la seguente nel caso di spazi vettoriali a dimensione finita.

Definizione 7.4.1 *Sia* $F : V \to V$ *un endomorfismo. Il numero reale* λ *è un autovalore di* F *se esiste* $v \in V$, $v \neq 0$, *tale che* $F(v) = \lambda v$.

Il vettore v *si dice autovettore di* F *relativo all'autovalore* λ.

Supporremo che tutti gli autovalori degli endomorfismi considerati siano reali (in generale appartenenti al campo base di V, cosa che è sempre verificata se V è un $\mathbb{C}$-spazio vettoriale).

Definizione 7.4.2 *Un sottospazio W di V si dice* invariante *per F (o anche F-invariante) se $F(W) \subseteq W$.*

Osservazione 7.4.3 *Se con $F_{|W}$ indichiamo la restrizione di F a W, cioè consideriamo solo le immagini di elementi di W, allora se W è un sottospazio F-invariante, $F_{|W} : W \to W$ è un endomorfismo di W.*

La scomposizione di V in somma diretta di spazi F-invarianti è un passo fondamentale della nostra ricerca.

Osservazione 7.4.4 *Supponiamo che $V = W_1 \oplus W_2$ con entrambi i sottospazi W_i F-invarianti. Sia $\mathcal{B}' = \{b'_1, \dots, b'_h\}$ una base di W_1 e $B'' = \{b''_1, \dots, b''_{n-h}\}$ una base di W_2. Chiaramente $\mathcal{B} = \mathcal{B}' \cup \mathcal{B}''$ è una base di V. Nella base $\mathcal{B}$ le coordinate di un vettore di W_1 saranno nulle a partire dal $(h+1)$-esimo posto, mentre saranno nulle le prime h coordinate di un vettore in W_2 (cfr. Proposizione 5.6.9).*

La matrice associata a F si può allora scrivere come matrice diagonale a blocchi:

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

con $A_1 \in \mathbf{M}_h(\mathbb{R})$ e $A_2 \in \mathbf{M}_{n-h}(\mathbb{R})$.

Infatti, per $i = 1, \dots, h$, la i-esima colonna di A è formata dalle coordinate di $F(b'_i)$ nella base $\mathcal{B}$. Poiché W_1 è F-invariante, $F(b'_i) \in W_1$ e quindi le sue ultime $n-h$ coordinate sono nulle.
Per $i = h+1, \dots, n$, la i-esima colonna di A è formata dalle coordinate di $F(b''_i)$ nella base $\mathcal{B}$. Poiché anche W_2 è F-invariante, $F(b''_i) \in W_2$ e quindi le prime h coordinate di $F(b''_i)$ sono nulle.

Osservazione 7.4.5 *La matrice A_1, per come è definita, esprime $F_{|W_1}$ nella base $\mathcal{B}'$ di W_1. Analogamente A_2 esprime $F_{|W_2}$ nella base $\mathcal{B}''$ di W_2.*

Proposizione 7.4.6 *Se $V = W_1 \oplus W_2$ con W_i sottospazi F-invarianti, allora il polinomio caratteristico di F si scrive come prodotto dei polinomi caratteristici di $F_1 = F_{|W_1}$ e $F_2 = F_{|W_2}$.*

Dim. Se $A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$, come nell'Osservazione 7.4.4, si ha:

$$\begin{aligned} p_F(t) &= p_A(t) = \det(A - tI_n) = \det \begin{pmatrix} A_1 - tI_h & 0 \\ 0 & A_2 - tI_{n-h} \end{pmatrix} \\ &= \det(A_1 - tI_h)\det(A_2 - tI_{n-h}) = p_{F_1}(t)p_{F_2}(t)\,, \end{aligned}$$

avendo utilizzato che il determinante di una matrice diagonale a blocchi è uguale al prodotto dei determinanti dei suoi blocchi. □

Osservazione 7.4.7 *Se λ è un autovalore di F_i, allora λ è anche un autovalore di F.*

Dedichiamo questa prima discussione alla ricerca di una scomposizione di V in sottospazi F-invarianti, dopo averne visto la utilità.

Osservazione 7.4.8 *Se λ è un autovalore di F, allora $V_1(\lambda) = ker(F - \lambda I)$, che è l'autospazio di F relativo a λ, è F-invariante.*

Se $x \in \mathrm{Ker}(F - \lambda I)$, cioè $(F - \lambda I)(x) = 0$, allora $F(x) \in \mathrm{Ker}(F - \lambda I)$. Infatti: $(F - \lambda I)(F(x)) = F^2(x) - \lambda F(x) = F((F - \lambda I)(x)) = F(0) = 0$.

Osservazione 7.4.9 *Gli endomorfismi F e $F - \lambda I$ commutano tra loro, cioè $F \circ (F - \lambda I) = (F - \lambda I) \circ F$, dove $\circ$ è la composizione dei due endomorfismi.*

Per brevità d'ora in poi useremo le notazioni:

$$T_\lambda = F - \lambda I\,, \tag{7.3}$$

$$V_k(\lambda) = \mathrm{Ker}(T_\lambda^k)\,. \tag{7.4}$$

Ad esempio, $V_2(\lambda) = \mathrm{Ker}(T_\lambda^2)$.

Osservazione 7.4.10 *Se $i < j$, allora $V_i(\lambda) \subseteq V_j(\lambda)$.*

Infatti, se $x \in V_i(\lambda)$, allora $T_\lambda^i(x) = 0$. Dunque $T_\lambda^j(x) = T_\lambda^{j-i}T_\lambda^i(x) = 0$ e quindi $x \in V_j(\lambda)$.

Abbiamo così costruito una catena ascendente di sottospazi di V.

$$0 \subseteq V_1(\lambda) \subseteq V_2(\lambda) \subseteq \ldots \subseteq V_n(\lambda) \subseteq \ldots \tag{7.5}$$

Poiché V è di dimensione finita, questa catena si dovrà stabilizzare, esisterà cioè un indice ν tale che: $V_{\nu-1}(\lambda) \subsetneq V_\nu(\lambda)$ e $V_k(\lambda) = V_\nu(\lambda)$ per ogni $k \geq \nu$. Indichiamo con $W(\lambda)$ tale $V_\nu(\lambda) = \mathrm{Ker}T_\lambda^\nu$.

Definizione 7.4.11 *Il sottospazio $W(\lambda)$ si dice* autospazio generalizzato *dell'autovalore λ.*

Osservazione 7.4.12 *Il sottospazio $W(\lambda)$ è F-invariante.*

La dimostrazione di ciò ricalca da vicino quella della Osservazione 7.4.8, poiché F, commutando con T_λ, commuta anche con T_λ^k per ogni k.

Problema 7.4.13 *Come determinare ν? Come posso essere sicuro che la catena (7.5) si è effettivamente stabilizzata?*

1) Sia i un indice tale che $V_i(\lambda) = V_{i+1}(\lambda)$. Allora si ha che $V_{i+1}(\lambda) = V_{i+2}(\lambda)$.

Basta far vedere che $V_{i+2}(\lambda) \subseteq V_{i+1}(\lambda)$, perché l'altra inclusione la conosciamo già. Sia $x \in V_{i+2}(\lambda)$, ossia $T_\lambda^{i+2}(x) = 0$. Cioè:

$$T_\lambda^{i+2}(x) = T_\lambda^{i+1} T_\lambda(x) = 0\,.$$

Da ciò si deduce che $T_\lambda(x) \in V_{i+1}(\lambda) = \mathrm{Ker} T_\lambda^{i+1}$. Per ipotesi, poiché $V_i(\lambda) = V^{i+1}(\lambda)$, $T_\lambda(x) \in V_i(\lambda) = \mathrm{Ker} T_\lambda^i$. Ossia $T_\lambda^i T_\lambda(x) = 0$. Ma allora $T_\lambda^{i+1}(x) = 0$, quindi $x \in V_{i+1}(\lambda)$.

2) Da ciò si deduce che la catena (7.5) è dapprima *strettamente* ascendente e poi si stabilizza definitivamente.

La (7.5) si riflette sulle dimensioni dei $V_i(\lambda)$. Si ha:

$$0 < \dim V_1(\lambda) < \ldots < \dim V_\nu(\lambda) = \dim V_{\nu+1}(\lambda) = \cdots \qquad (7.6)$$

Siccome $\dim V_i(\lambda) = n - \mathrm{car}(T_\lambda^i)$, la catena diventa:

$$n > \mathrm{car}(T_\lambda) > \ldots > \mathrm{car}(T_\lambda^\nu) = \mathrm{car}(T_\lambda^{\nu+1}) = \cdots$$

Quindi, operativamente, calcoliamo $\mathrm{car}(T_\lambda^i)$. Il primo indice ν per cui due caratteristiche consecutive sono uguali, è, per 1), l'indice per cui la catena (7.5) si stabilizza e che ci permette di determinare $W(\lambda)$, l'autospazio generalizzato relativo all'autovalore λ, essendo $W(\lambda) = V_\nu(\lambda)$.

A questo punto enunciamo il risultato più interessante di questa sezione.

Teorema 7.4.14 *Sia $\{\lambda_1, \ldots, \lambda_m\}$ l'insieme degli autovalori dell'endomorfismo F di V. Se $W(\lambda_i)$ è l'autospazio generalizzato relativo a λ_i, si ha:*

$$V = W(\lambda_1) \oplus \cdots \oplus W(\lambda_m)\,. \qquad (7.7)$$

Posponiamo la sua dimostrazione alla esposizione di alcune osservazioni che sarà utile tener presenti.

Osservazione 7.4.15 *1)* $T^k_{\lambda_i}$ *e* $T^h_{\lambda_j}$ *commutano.*
(Ricordiamo che due endomorfismi F *e* G *commutano se* $F \circ G = G \circ F$*).*

2) $F_{|W(\lambda)}$ *è un endomorfismo di* $W(\lambda)$ *che ha come unico autovalore* λ*.*

3) Se $i \neq j$ *,* $W(\lambda_i) \cap W(\lambda_j) = \{0\}$*.*

4) $V = W(\lambda_i) \oplus \mathrm{Im} T^{\nu_i}_{\lambda_i}$*, dove* ν_i *è tale che* $W(\lambda_i) = V_{\nu_i}(\lambda_i)$*.*

5) $W(\lambda_i)$ *e* $\mathrm{Im} T^{\nu_i}_{\lambda_i}$ *sono sottospazi* F*-invarianti e anche* T_{λ_j}*-invarianti per ogni autovalore* λ_j *di* F*.*

6) Se $i \neq j$*,* $\mathrm{Ker}(T^h_{\lambda_j}) = V_h(\lambda_j) \subseteq \mathrm{Im} T^{\nu_i}_{\lambda_i}$*.*

La dimostrazione di queste osservazioni appesantirebbe il discorso facendo perdere il filo logico della costruzione (ammesso che non lo si sia già perso). Chi volesse, la può trovare alla fine di questo capitolo.

Dim. (del Teorema 7.4.14). Dimostriamo questo teorema per induzione su $n = \dim V$.

Se $\dim V = 1$, allora $V = W(\lambda)$, perché sempre $\dim W(\lambda) > 0$.

Supponiamo che l'enunciato sia vero per tutti gli spazi vettoriali aventi dimensione minore di n, per endomorfismi che abbiano tutti gli autovalori reali. Dimostriamolo per V di dinensione n.

Sia λ_1 un autovalore di F. Dalla Osservazione 7.4.15 (punto 4) abbiamo che:

$$\begin{aligned} &V = W \oplus U, \quad \text{dove } W = W(\lambda_1) \text{ e } U = \mathrm{Im} T_{\lambda_1} \nu_1, \\ &\dim U < n\,, \quad \text{perché } \dim W \geq 1. \end{aligned}$$

Poiché W e U sono F-invarianti, per la Proposizione 7.4.6 il polinomio caratteristico di F, $p_F(t)$, si scrive come prodotto:

$$p_F(t) = p_{F_{|W}}(t) p_{F_{|U}}(t)\,.$$

Per l'Osservazione 7.4.15 (punto 2), $F_{|W}$ ha come unico autovalore λ_1. Allora tutti gli altri autovalori di F, $\lambda_2, \ldots, \lambda_m$, che sono radici di $p_F(t)$, debbono essere radici di $p_{F_{|U}}(t)$ e quindi λ_i, per $i = 2, \ldots, m$, sono tutti autovalori

di $F_{|U}$. $F_{|U}$ non può avere autovalori che non siano anche autovalori di F, pertanto per U scatta l'ipotesi induttiva:

$$U = \hat{W}(\lambda_2) \oplus \cdots \oplus \hat{W}(\lambda_m)\,,$$

dove $W = \mathrm{Ker}((T_{\lambda_i})^{\mu_i}_{|U})$ e μ_i è un opportuno esponente.

Se risultasse che:

$$\mathrm{Ker}((T_{\lambda_i})^h_{|U}) = \mathrm{Ker}(T^h_{\lambda_i}) \quad \text{per ogni esponente } h\,, \tag{7.8}$$

allora $W = W(\lambda_i)$ e il teorema sarebbe dimostrato.

Ma, per l'Osservazione 7.4.15 (punto 6), $\mathrm{Ker}(T^h_{\lambda_i}) \subseteq U$ per ogni esponente h e allora l'uguaglianza (7.8) è banalmente vera. □

Una conseguenza della scomposizione di V in sottospazi invarianti:

$$V = W(\lambda_1) \oplus \cdots \oplus W(\lambda_m)\,,$$

è che il polinomio caratteristico di F risulta scomposto in fattori ciascuno dei quali è il polinomio caratteristico di $F_i = F_{|W(\lambda_i)}$:

$$p_F(t) = p_{F_1}(t) \cdots p_{F_m}(t)\,.$$

Poiché, per l'Osservazione 7.4.15 (punto 2), λ_i è l'unico autovalore di F_i, $p_{F_i}(t) = (\lambda_i - t)^{\alpha_i}$. Poiché il grado del polinomio caratteristico è uguale alla dimensione dello spazio, $\alpha_i = \dim W(\lambda_i)$.

Abbiamo la seguente scomposizione di $p_F(t)$:

$$p_F(t) = (\lambda_1 - t)^{\alpha_1} \cdots (\lambda_m - t)^{\alpha_m}\,.$$

Per l'unicità della scomposizione di un polinomio in fattori di primo grado, si ha che $\alpha_i = \dim W(\lambda_i) = m_a(\lambda_i)$, la molteplicità algebrica di λ_i.
In particolare, tenendo presente che $\mathrm{Aut}(\lambda_i) \subseteq W(\lambda_i)$, si ha:

$$1 \leq m_g(\lambda_i) = \dim \mathrm{Aut}(\lambda_i) \leq \dim W(\lambda_i) = m_a(\lambda_i)\,. \tag{7.9}$$

Inoltre, se scelgo una base di V che sia l'unione delle basi dei singoli $W(\lambda_i)$, la matrice associata a F sarà una matrice a blocchi del tipo:

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & A_m \end{pmatrix}\,,$$

dove $A_i \in \mathbf{M}_{\alpha_i}(\mathbb{R})$ è la matrice associata a F_i nella base scelta di $W(\lambda_i)$.

Occupiamoci adesso di determinare la base più opportuna in ciascuno dei $W(\lambda_i)$ per l'endomorfismo F_i (che, ricordiamo, ha un solo autovalore), in modo da rendere ciascuna matrice A_i più semplice che sia possibile.

Definizione 7.4.16 *Un endomorfismo $T : V \to V$ si dice* nilpotente *se esiste un intero positivo k tale che $T^k = 0$.*
Il minimo valore m di k per cui ciò accade si dice ordine di nilpotenza *di T.*

Osservazione 7.4.17 *Se F è un endomorfismo con un solo autovalore λ, allora $F - \lambda I$ è un endomorfismo nilpotente.*

Infatti, $V = W(\lambda)$, per il teorema di scomposizione. Ma $V = W(\lambda) =$ $\mathrm{Ker}(F - \lambda I)^\nu = 0$ *e quindi $(F - \lambda I)^\nu = 0$. Dunque tutti gli endomorfismi F_i sono tali che gli endomorfismi $(F_i - \lambda_i I)$ sono nilpotenti.*

Se in una base opportuna un endomorfismo F è espresso dalla matrice A_F, nella stessa base, $T = F - \lambda I$ è espresso dalla matrice $A_T = A_F - \lambda I_n$, dove I_n è la matrice identità di ordine n. Quindi una base che rende semplice la matrice di $F - \lambda I$ rende semplice anche la matrice di F.

Cerchiamo allora matrici semplici per endomorfismi nilpotenti.

Proposizione 7.4.18 *Sia T un endomorfismo nilpotente di ordine m. Allora esiste $v \in V$ tale che $T^{m-1}(v) \neq 0$. Inoltre gli elementi di V:*

$$v, T(v), T^2(v), \ldots, T^{m-1}(v),$$

sono linearmente indipendenti.

Dim. Poiché l'ordine di nilpotenza è m, $T^{m-1} \neq 0$, quindi esiste $v \in V$ tale che $T^{m-1}(v) \neq 0$.

Supponiamo che sia:

$$\mu_0 v + \mu_1 T(v) + \cdots + \mu_{m-1} T^{m-1}(v) = 0\,.$$

Sia h il minimo intero tale che $\mu_h \neq 0$. Allora:

$$0 = T^{m-h-1}(\mu_h T^h(v) + \cdots + \mu_{m-1} T^{m-1}(v)) = \mu_h T^{m-1}(v) = 0\,.$$

Ma ciò è assurdo perché sia μ_h che $T^{m-1}(v)$ sono diversi da zero. □